COGNITIVE AND COMPUTATIONAL MODELING OF HANDWRITING

by

Gali Ellenblum

A dissertation submitted to Johns Hopkins University in conformity with the requirements

for the degree of Doctor of Philosophy

Baltimore, Maryland

February, 2019

Abstract

Handwritten letters can be produced in many different ways – there is nothing inherent in the shape of a letter that dictates in which order and direction the strokes ought to be produced. For example, many people write uppercase ‘A’ starting at the top-left, while others choose to start it at the bottom-left. But uppercase ‘A’ could also be written starting with the horizontal line, yet this stroke pattern is never observed. Why are some stroke patterns observed while others are not? And how is one pattern chosen over others in a given instance? The systematicity in the observed productions of letters has brought researchers in the past (e.g., Goodman & Levine, 1973; Van Sommers, 1984), to propose that rules govern the way letters are written. But the rules that have been proposed, for example ‘start at the top’ or ‘no pen lifts,’ are often violated and occasionally come into conflict. In this dissertation, we present an account that can deal with these conflicts and with rule violations, using a novel application of Optimality Theory (OT; Prince & Smolensky, 1993). At the center of OT is the notion that rules can be violated, and that conflicts between them are resolved by a ranking of the rules. In addition to OT, we also implemented Harmonic Grammar (HG; Legendre, Miyata, & Smolensky, 1990a), a sister-framework to OT that requires only a weighting of the rules, rather than a ranking. We develop the theoretical framework needed to apply OT and HG to handwriting, and show that we can use these systems to model both prescribed and participant writing, in English and in Hebrew, by right-handed and left-handed individuals. We further use the framework of violable rules to resolve previously unsettled debates, and to shed light on some of the cognitive mechanisms underlying handwriting.

Dissertation committee

Michael McCloskey (primary advisor), Cognitive Science

Paul Smolensky, Cognitive Science

Colin Wilson, Cognitive Science

Nazbanou Nozari, Neurology

Jonathan Flombaum, Psychological and Brain Sciences

Acknowledgments

I am extremely grateful to Mike McCloskey for the mentorship and support he has given me throughout this entire process. I am also thankful to Paul Smolensky, Colin Wilson, and Brenda Rapp for giving me guidance and advice through the writing of this dissertation that undoubtedly made the finished product better, and to Bonnie Nozari and Jon Flombaum for their insightful comments and questions on the dissertation.

I would like to thank all the research assistants with whom I have worked through the years, who collected and analyzed data, and especially Kim Wong, who was my right-hand and go-to person with any analyses for which I needed someone I can trust completely. Thank you to the entire McCloskey Lab for fruitful discussions of this work, and for coming to an undisclosed number of practice-talks before my defense. Thank you to all the graduate students in the Cognitive Science Department. To Jeongho and Pang who were the other edges of the best TA trio in history, and to Bob Wiley, who was a friend throughout this experience.

Thank you to my parents, Ronnie and Lenore, my sister Maya and my brother Yuval, for journeying across the Atlantic to see us, and for always supporting me from afar. And to my parents- and sisters-in-law for helping Dan shoulder the burden when I couldn't. Thank you to Roi for always being there for me.

This work is dedicated to my better half, Dan, without whom I could not have done this, and to the light of my life, Ori, who taught me that great things are worth waiting for.

I wouldn't trade my scars for the lessons I've learned.

Gali

Table of contents

List of tables

List of figures

Introduction

How are letters written? Are the pen movements that people use to produce letter-shapes arbitrary, or do they adhere to some organizing principles? What are those principles, and how are they structured? And what is the nature of the relationship between the principles? These are some of the questions we seek to answer in this dissertation. We use a combination of behavioral and computational methods to explore the cognitive mechanisms underlying handwriting in different languages, by different individuals, and using different effectors.

Cognitive theories of spelling and writing

Theories of orthographic processing often distinguish between the cognitive mechanisms required for reading and those required for writing and spelling. Although reading tends to receive more attention, writing may be an equally crucial function for everyday life. In a study involving stroke survivors, 71% of participants reported writing and spelling impairments being in the top 5 most important problems they have had to face since their stroke (Hillis & Tippett, 2014). According to the study, writing and spelling impairments were more important than impairments of reading (50%), word retrieval (43%), walking (50%), or motor function in general (57%).

Models of writing and spelling are usually in agreement that the process of writing involves several separate components. Flower and Hayes (1981) describe three stages in the production of written output: Planning, translating, and reviewing. The planning stage takes ideas from abstract concepts into organized thoughts to be written, the translating stage turns those thoughts into visible language, and in the reviewing stage writers evaluate and revise what they wrote to better reflect the meaning they intended to convey. The vast majority of research on writing and spelling has focused on the processes roughly equivalent to the “translating” stage in Flower and Hayes’ Model (1981). The Dual Route Model (DRM; e.g., Ellis, 1988, 1989; Rapp & Caramazza, 1997; Tainturier & Rapp, 2001) is perhaps the most widely accepted framework for the study of writing and spelling, as for the study of reading.

According to the DRM (Figure 1), a word to be produced is chosen from some input (either auditory input such as a dictated word, or internal conceptual input such as a person's memory or imagination). The word is then processed through one of two separate routes: The lexical route, or the sub-lexical route. In the lexical route, the spelling of the word is retrieved from a long-term memory storage unit known as a lexicon. In the sub-lexical route, the spelling is constructed by breaking the word down into its phonemes, and directly converting the phonemes to appropriate graphemes. The lexical route is useful for spelling words that the speller has encountered before (i.e., words whose spelling is known), and especially words with irregular grapheme-phoneme correspondence (e.g., "gauge"). The sub-lexical route, which may construct a wrong spelling based on grapheme-phoneme correspondences, is used to spell words whose spelling is not found in the lexicon, such as non-words (e.g., "yoap") or words the speller has never seen before (Tainturier & Rapp, 2001). Once the spelling for a particular word has been determined, a short-term memory component called the graphemic buffer is responsible for keeping active the letter identities corresponding to the correct letters to be produced. The graphemic buffer then releases those letters in the correct order so that they can be produced through oral spelling, handwriting, or typing (Buchwald & Rapp, 2004).

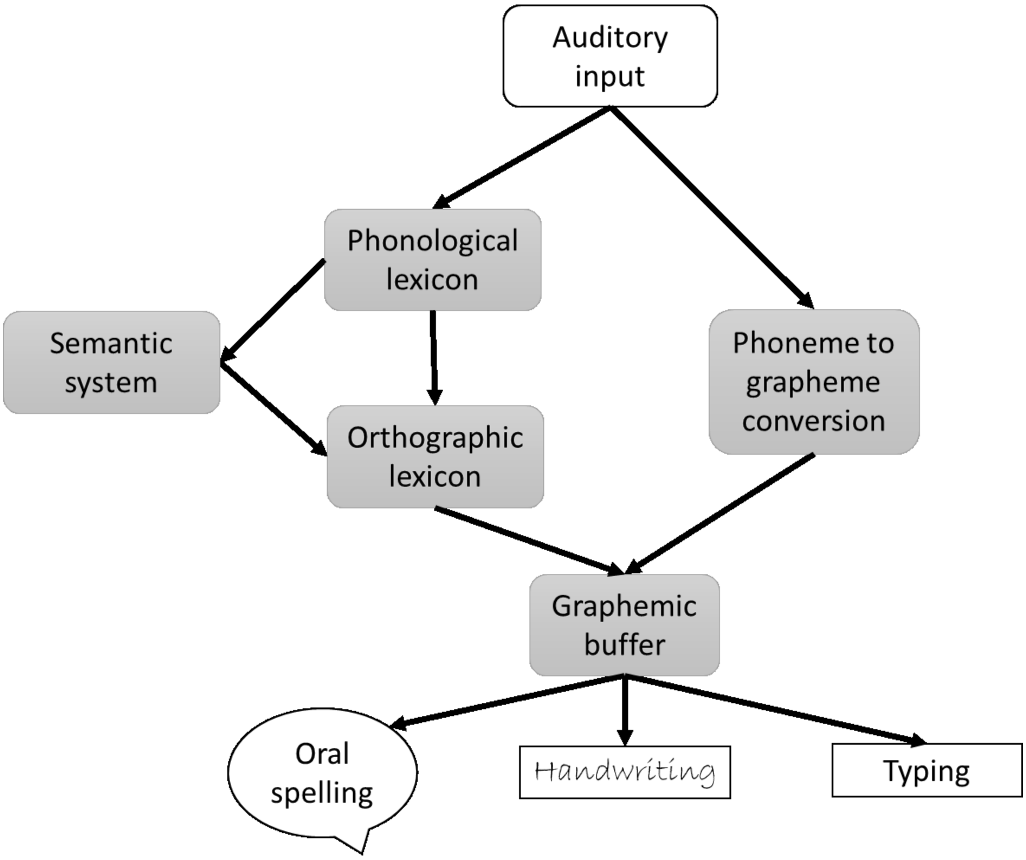


Figure 1. A schematic of the dual-route model of spelling and writing.

Most of the research concerning writing has focused on the processes involved in spelling, whether “central” (i.e., the lexical and sub-lexical routes), or “peripheral” ones (i.e., involving the mechanisms operating before or after the two routes diverge; e.g., Caramazza & Miceli, 1990; Purcell, Jiang, & Eden, 2017; Purcell, Turkeltaub, Eden, & Rapp, 2011; Rapp, Purcell, Hillis, Capasso, & Miceli, 2016). Studies of spelling often employ neuropsychological methods, and attempt to characterize the function of each of the components of the spelling model using acquired or developmental impairments (e.g., Hillis & Caramazza, 1991; Fischer-Baum, McCloskey, & Rapp, 2010; McCloskey, Badecker, Goodman-Shulman, & Aliminosa, 1994; Miozzo & De Bastiani, 2010; and see McCloskey & Rapp, 2017 for a review of developmental dysgraphias).

A lot less has been written on the processes following the selection and ordering of graphemes to be produced. To produce handwritten characters, the letters stored in the graphemic buffer, which at that point are only abstract identity representations (i.e., they do not include information about the letter shape, name, or sound), need to trigger the initiation of a series of motor movements to produce the desired shape (e.g., Ellis, 1988; Margolin, 1984; Rapp & Caramazza, 1997). Most researchers identify three processes between the graphemic buffer and the output shapes (e.g., Ellis, 1988; Margolin, 1984). First, the particular allograph needs to be chosen for the letter in question (for example, lowercase a or uppercase A, or z with a crossing line or without). Next, effector-independent motor plans matching the allograph in question are activated. The effector independent motor plans specify only the direction, shape, and relative location the pen-movements should take, regardless of the effector used to produce them. Finally, the effector-independent motor plans are converted into effector-specific motor programs, which specify the exact muscles and joints needed to produce the movements, and those are then executed by the chosen effector. While this three-step model is common, note that some researchers have proposed that the abstract letter identities are mapped directly onto the effector-independent motor plans, without passing through the allographic representation (e.g., Rapp & Caramazza, 1997).

In recent years, several groups have described the importance of handwriting in facilitating reading processes (e.g., Berninger, Abbott, Abbott, Graham & Richards, 2002; James & Engelhardt, 2012; James & Gauthier, 2011; Kersey & James, 2013; Wiley, Wilson, & Rapp, 2016; Wong, Wadee, Ellenblum, & McCloskey, 2018). Other researchers have attempted to characterize the relationship between handwriting and the acquisition of spelling and motor processes (e.g., Kandel, Peereman, & Ghimenton, 2013; Kandel & Perret, 2015). More recently, McCloskey, Reilhac, and Schubert (2018) described a case of an acquired dysgraphia patient who had intact oral spelling but impaired written spelling, noting that the patient's errors in writing seemed to stem not from the selection of a wrong letter identity or shape to be produced, but rather from a degraded representation of the motor plans needed to produce the shapes (McCloskey et al., 2018). A somewhat related paper by Law, Ki, Chung, Ko and Lam (1998) described the way Chinese children produce Chinese characters, and how they diverge from the prescribed writing in Chinese.

A group of researchers led by Brenden Lake (Lake, Salakhutdinov, Gross, & Tenenbaum, 2011; Lake, Salakhutdinov, & Tenenbaum, 2012, 2013, 2015) has also modeled writing production, but with a slightly different perspective and with different goals than what we intend to do. Our goal was to understand why people choose to produce character-shapes a certain way, and what we can learn about the rules governing the way characters are produced. Lake et al. (2011, 2012, 2013, 2015), on the other hand, used the production of characters to learn about the internal representation of the character in terms of the motor patterns involved in producing it. They then introduce a classifier that uses a similar structure of representation to the one they have identified in human participants, to perform "one-shot learning" similar to what people do when they classify newly encountered characters (Lake et al., 2011, 2012, 2013, 2015). One major difference between our endeavor and Lake et al.'s, is that while they are searching for the invariant aspects of character production, we are particularly interested in the variability observed in the production of known shapes, between different people, different languages, and other differentiating factors.

Systematicity and variability in written letter production

Writing letters seems straightforward: You move the pen to create the shape you have known since first grade. But many decisions are involved in determining how to produce the shape. The production of letters can be described in many ways, and one can choose to model handwriting at different levels. On some level of analysis, characters can be broken down into strokes. For example, the letter 'A' can be described as two diagonal strokes and one horizontal stroke. We will assume that letters have an underlying representation consisting of those strokes (in accord with previous research such as Edelman & Flash, 1987, and Kandel & Spinelli, 2010, but see Hollerbach, 1980 for a model that is not based on this kind of representation). However, there is nothing inherent to the shape of a letter that dictates the order or direction of production of the strokes needed to produce it. Indeed, the strokes comprising a letter can be written many different ways, creating different stroke patterns depending on the direction of each of the strokes (e.g., producing the left-diagonal line in ‘A’ from the top down or from the bottom up), and their order (e.g., starting with the left diagonal or the right diagonal stroke in ‘A’).

And yet, most possible stroke patterns for each letter are never observed. Why are some stroke patterns chosen over others? And why are some never produced? In this dissertation, we lay out a framework that attempts to answer these questions and deal with some of the vagaries of handwriting. To this end, we recruit Optimality Theory (OT, Prince & Smolensky, 1993), a framework suitable to deal with hierarchical rule-systems in which the rules often conflict and are violated. We seek to find whether a series of ranked, violable rules (or constraints) can account for stroke patterns in handwriting of letters, in prescribed or actual writing, in different languages, and by right and left-handed individuals. We also recruit an implementation of Harmonic Grammar (HG, Legendre, Miyata, & Smolensky, 1990a), a sister-approach to OT that similarly deals with complex relationships between different rules. If handwriting can indeed be modeled using OT or HG, we want to know what else we can learn from the modeling endeavor, and whether we can answer other questions about the cognitive underpinnings of handwriting within this framework. To achieve these goals, we must begin with a simple question: Why do we write letters the way we do?

Both casual observation and previous studies reveal significant variability in handwritten production of letters and other shapes by different people (e.g., Thomassen, Meulenbroek, & Tibosch, 1991; Wing, 1979). Consider for example the three strokes of the letter ‘A’ (Figure 2). This shape can be produced in several different ways: Many people would write it starting at the top, and going down the left diagonal stroke first, then down the right one, and finally crossing the horizontal stroke from left-to-right (Figure 2, left). Others might start the left stroke at the bottom, and go up and then down the right stroke (Figure 2, center). While both of these stroke patterns are very common, there are many other possible patterns that are almost never produced. For example, it would seem extremely odd and unusual for English speakers to begin writing the letter ‘A’ with the horizontal stroke and then produce the right diagonal line before the left one (Figure 2, right). So why is it the case that some patterns are observed while others are not?

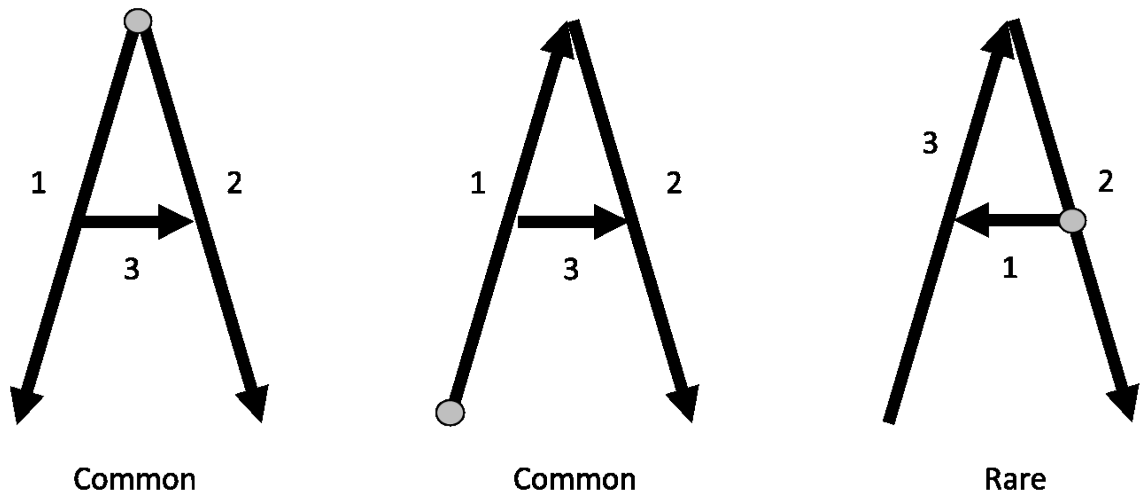


Figure 2. Three possible ways to write uppercase ‘A’. The grey dot marks the starting point, arrows mark the direction of production, and the numbers mark the order of production.

The stroke patterns used to produce written characters do not seem to be chosen arbitrarily. Instead, they seem to follow some organizing principles. Recognizing these underlying principles, researchers (e.g., Goodnow & Levine, 1973; Van Sommers, 1984) have proposed rule systems that would explain why some stroke patterns are common while others are rare. Goodnow and Levine (1973) termed this rule system “the Grammar of Action”, suggesting that children from pre-school to first grade, as well as adults, follow certain rules when producing shapes and characters in writing. They proposed seven major rules, including ‘start at the top’, ‘draw horizontals from left-to-right,’ and ‘thread if possible’ (i.e., continue to the next stroke without lifting the pen, if the two strokes are immediately adjacent). They observed that there are hierarchies among these rules, some of the rules being more important than others, and that some of those hierarchies seem to change with development. They also noted that children make more errors on shapes in which some rules conflict.

Building on Goodnow and Levine's (1973) work, several other groups have described rule systems governing handwriting. For example, Ninio and Lieblich (1976) have suggested rules of anchoring – starting a new stroke from a point that has already been drawn on the paper – and have shown how these occasionally conflict with the rule dictating left-to-right production of horizontal strokes. Nihei (1983) suggested a developmental shift, coinciding with the beginning of formal training in writing, in the relative importance given to different rules: From the preference of children to anchor movement to that of adults to produce horizontal strokes from left-to-right.

The most thorough description to date of the principles and rules governing handwriting can be found in 'Drawing and Cognition' (Van Sommers, 1984). In this book, Van Sommers gives an extensive account of handwriting and drawing, accompanied by numerous experiments conducted with children and with adults. He discusses many of the underlying mechanisms governing handwritten production of shapes in copying and in free-form drawing by right-handed and left-handed people of multiple cultural and language backgrounds. Among other things, he discusses some of the mechanisms dictating stroke direction, such as good control over the motion of the effector holding the writing instrument (e.g., good control of hand-movement when holding a pen), adherence to language-specific directionality, and preventing obstruction of view of the target to be copied. He dedicates an entire chapter of his book to the production of curves and circles, observing that the direction in which a circle is produced depends on handedness, native language, and origin point on the circumference. Further relevant to our enterprise is Van Sommers's (1984) discussion of threading, and the occasions on which a conflict between threading and preferred stroke direction results in one of these rules having the upper hand.

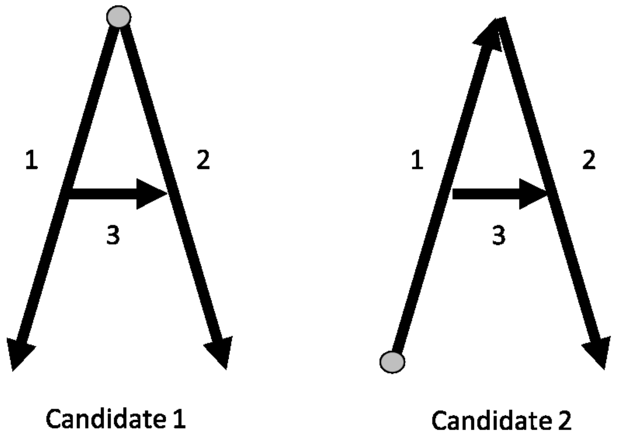
While Van Sommers (1984), Goodnow and Levine (1973), and others have come a long way in describing rules and principles that govern handwritten production, none of them have suggested a systematic framework of rules that can account for all the letters in a script. Goodnow and Levine (1973) have identified some rules that account for particular choices that English-speaking children and adults make when writing (e.g., writers generally start with a vertical line, and progress from left-to-right). However, those rules only account for most of the participants and for some of the shapes, and they are unable to explain cases in which production deviates from the stated rule. While Goodnow and Levine (1973) acknowledge that some rules conflict (e.g., when starting with a vertical stroke necessitates progression from right-to-left), they do not offer a comprehensive account of how these conflicts are resolved. Other researchers, such as Nihei (1983) have focused on developmental changes in the application of some rules, and in the way that conflicts are resolved (e.g., when there is a conflict, children aged 4-5 years prefer to anchor their strokes, whereas children aged 5-6 years prefer to thread them). A discussion of systematicity stemming from cultural or language-specific characteristics (e.g., Goodnow, Friedman, Bernbaum, & Brauch Lehman, 1973), showing that progression between strokes often follows the overall direction of reading, nevertheless stops short of a thorough analysis of cases in which the overall reading direction conflicts with the direction of best controlled movement (e.g., in left-handed English speakers).

A conflict between rules can manifest in three major ways: (1) Within letters – when the production of a character requires prioritizing one rule over another; (2) across a set of letters – when the production of one character is shown to follow a particular rule, which is then violated on the production of another character; and (3) across participants – when a rule followed by one person producing a certain character is violated by another person producing the same character. None of the above-mentioned studies propose a framework that can account for those three types of conflict. Instead, researchers so far have dealt with these conflicts by showing a general tendency towards some rule over another by most people. For example, Goodnow and Levine (1973) have shown that approximately 80% of English-speaking adults preferred to produce all vertical lines from top to bottom; Van Sommers (1984) has shown that most right-handed English-speaking adults preferred to produce circles counter-clockwise, etc. However, both Goodnow and Levine (1973) and Van Sommers (1984) recognize that their proposals do not fully account for observed stroke patterns (e.g., “accounting for [order in successive behavior] is recognized as a continuing problem in psychology”, Goodnow & Levine, 1973, p. 94).

And therein lies the rub: The fact that most possible stroke patterns for letters are never observed suggests that there are rules governing the production, but letter production necessarily entails conflicts and violations of the rules that have been proposed thus far. We attempt to reconcile the two by building on the notion of rules that are grounded in basic principles of motor control, reading direction, legibility, speed, etc., and accepting that some of the rules that have been suggested previously are correct. The crux of our approach is dealing with the conflicts between these rules or principles – within letters, across letters, and across individuals. It does not seem that handwriting follows the simplest kind of rule system, where all the rules always apply. Instead, it follows a complex system involving “softer” rules – or constraints, that can be violated, and are organized in a hierarchical system in which some constraints are more important than others. To deal with this complex system, we adopt the framework of Optimality Theory (OT, Prince & Smolensky, 1993).

Optimality Theory

OT (Prince & Smolensky, 1993), which was first developed in the context of phonology, has been very influential in various fields, mainly in phonology and syntax (Boersma, Dekkers, & Weijer, 2000), and has also been used outside the realm of traditional linguistics, for example to account for the presence of different shapes in character inventories (Primus, 2004). One of the advantages of OT for our purposes is that OT acknowledges that it is not always possible to satisfy all the rules at the same time, and aims to solve the problem of selecting an optimal candidate among many competing candidates in the context of violated and conflicting rules. In our case, those competing candidates are different ways of writing the strokes comprising a character (or different stroke patterns). OT proposes that universal grammar consists of a set of constraints, which unlike hard-set rules, are violable. While any constraint could be violated any number of times, not all constraints are equal in their importance. The constraints are ranked relatively to one another, so that the violation of a higher-ranked constraint incurs a more severe penalty than the violation of lower-ranked constraints. OT uses a special kind of weighting called strict domination, in which one violation of a higher-ranked constraint is worse than any number of violations of lower-ranked constraints. We demonstrate this principle with an example.



Constraint	Candidate	1	2
	[redacted]		
	No pen lifts	2	1

Figure 3. Left: Two possible ways to write uppercase 'A'; Candidate 1 (left) starts at the top and proceeds down the left line, and Candidate 2 starts at the bottom left and travels up. The grey dot marks the starting point; the number marks the order of production. Right: Profile of constraint violations for the two Candidates.

Let us consider a simplified system with only two constraints: 'start at the top' and 'no pen lifts,' and two possible stroke patterns for the letter 'A', which we will call candidates, illustrated in Figure 3. While both candidates are reasonable ways to write the letter 'A', each candidate violates a constraint: Candidate 1 violates the constraint 'no pen lifts' twice (when transitioning between the first and second strokes, and again when transitioning between the second and third strokes), whereas candidate 2 violates this constraint only once (there is no pen lift when transitioning between the first and second strokes). However, candidate 2 also violates the constraint 'start at the top,' as the first stroke starts at the bottom and moves upwards, which candidate 1 does not violate.

The "optimal candidate" according to OT is determined by the principle of strict domination: One violation of a higher-ranked constraint is worse than any number of violations of lower ranked constraints. Therefore, in our simplified example, if ‘start at the top’ is more important, then candidate 1 “wins” (i.e., is the optimal candidate), since it violates this constraint fewer times than candidate 2. However, if ‘no pen lifts’ is more important, then candidate 2 wins, because it violates this constraint fewer times than candidate 1 does.

It follows from the explanation above that to use OT to determine the optimal candidate we need two things: An inventory of character shapes, from which we can derive all possible candidates, and a list of constraints from which to compute constraint violations. While defining the character shapes is relatively straight-forward, choosing and defining the constraints to be used is not a trivial task. We want the constraints to be well-motivated, and to follow general principles that are important to handwriting (such as maintaining legibility, speed, etc.) as much as possible. We also want the constraints to be detailed enough that it is clear how they apply in each case, and broad enough that each constraint applies to more than just one character or a handful of candidates. We will discuss the selection of constraints in more detail in Chapter 1.

If we have a set of constraints, and we know their ranking (i.e., we know the relative importance of the constraints), the optimal candidate for a given letter is easy to pick out (as in the example in Figure 3). But the ranking is not explicitly available, even once we have established a list of constraints, and needs to be determined. Using OT, we can infer the ranking of the constraints by observing the candidate that is produced among all the candidate stroke patterns for a given letter. Crucially, however, a ranking will only be found if all the stroke patterns produced (e.g., for all the letters in a language) are consistent in their adherence to a hierarchy of constraints (i.e., they all follow the same ranking of constraints). In the simplified example above (Figure 3), assuming there are only two candidates and two constraints, if the production we observe is of candidate 1, this means that ‘start at the top’ is more important than ‘no pen lifts’. On the other hand, if the production we observe is of candidate 2, this means that ‘no pen lifts’ is more important.

Even if our set only contained one letter, reaching a successful ranking that chooses one stroke pattern over all the others would not be trivial. A successful ranking means that the whole enterprise, using a ranking system to determine the optimal candidate, is viable, at least on a small scale. Of course, seeing as our set of candidates consists of more than one letter, the complexity of the problem increases. The set we are trying to model consists of an entire character inventory for a given language and script. In modeling all the letters together, one must consider dependencies in the ranking between letters. One way to think about it is that each character creates a hypothesis for other characters. We are attempting to find one ranking of constraints that would consistently pick the correct candidate in each of the characters in a language.

Assume for example that we observed the production of ‘A’ starting at the top (Figure 4, left), in a simplified system with only 2 constraints (‘start at the top’ and ‘no pen lifts’) as described earlier. Now consider the letter ‘N’: If the ranking of our simplified 2-constraint set is consistent across letters, and assuming no other constraints are at play, we should expect the production of ‘N’ using the stroke pattern designated candidate 1 (Figure 4, middle). If we observe instead candidate 2 (Figure 4, right), which starts at the bottom but does not lift the pen, then the ranking of the two constraints we have is inconsistent across our character system. Of course, we may be able to find a third constraint that is ranked higher than our two initial constraints and could help explain the observed production of candidate 2 for N (Figure 4, right), but again, the new ranking would need to apply to all other letters in the language. In short, our goal is to find a consistent ranking that would account for all the letters in a set. Since stroke patterns often violate different constraints, such a ranking is not trivial to find.

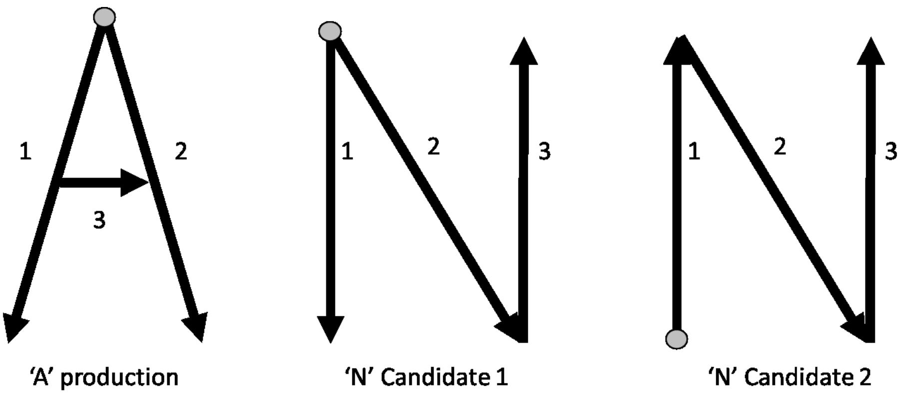


Figure 4. Left: A possible way to write uppercase 'A'; starting at the top and proceeding down the left line. Middle and right: Two possible ways to write the uppercase 'N': Candidate 1 starts at the top and lifts the pen once, whereas candidate 2 starts at the bottom and does not lift the pen. The grey dot marks the starting point; the number marks the order of production.

Harmonic Grammar

Harmonic grammar (HG), proposed by Legendre, Miyata, and Smolensky (1990a), is another formal framework, closely related to OT, that deals with violable constraints. Unlike OT, in which constraints are ranked in strict domination, HG selects the most well-formed candidate (or the most Harmonic output) based on a weighted sum of constraint violations (Potts, Pater, Jesney, Bhatt, & Becker, 2010). Instead of strict domination, a candidate's Harmony in HG is computed by penalizing each constraint violation by the weight for that constraint, and then summing the penalties over all the constraints. Therefore, whereas in OT a violation of a higher-ranked constraint is worse than any number of violations of lower-ranked constraints, in HG a candidate violating a higher-weighted constraint might still win over a candidate that violates only lower-weighted constraints if the lower-weighted constraints are violated enough times.

Consider for example the hypothetical tableau of constraint violations below (Table 1). Under strict domination (as in OT), if constraint 1 is ranked higher than constraint 2 then candidate 1 wins, because it violates constraint 1 fewer times. However, if we assume that constraint 1 has a weight of

3 and constraint 2 a weight of 2, then under HG candidate 1 has a harmony of -6 (it was penalized three times for violating constraint 2, and those penalties were subtracted from 0), and candidate 2 a harmony of -5 (it was penalized once for violating constraint 1 and once for violating constraint 2). Thus, even though it violates a higher-ranked constraint more times than candidate 1, candidate 2 is still more Harmonic, and would be chosen as the winner. As in OT, one can ask whether there is a set of weights that would choose all and only target candidates in HG. A set of weights that satisfy this demand can be achieved with a standard supervised learning algorithm.

	WEIGHT	**CANDIDATE 1**	**CANDIDATE 2**
CONSTRAINT 1	3	0	1
CONSTRAINT 2	2	3	1

Table 1. Hypothetical tableau indicating the number of constraint violations in a system with 2 constraints and 2 candidates. The weight associated with constraint 1 is 3 and the weight associated with constraint 2 is 2.

HG has been used in the past as a framework for studying linguistic systems (e.g., Legendre, Miyata, & Smolensky, 1990a, 1990b; Pater, 2009). Legendre et al. (1990a) suggest that the final steps in HG analysis should be "to interpret [the Harmony function] as embodying soft grammatical and lexical rules, [and to] analyze these rules for new linguistic insights into the original linguistic problem" (p. 10). We are using HG here because it allows us to answer some of the cognitive questions posed in this paper more fully and easily. For example, in cases where we fail to find an OT ranking that accounts for the data, we ask whether it is because there are inherent inconsistencies in people's handwriting, or whether, instead, the nature of strict domination is such that it does not fully capture how people produce stroke patterns. Furthermore, while OT is useful in determining whether or not a system of stroke patterns is entirely consistent (i.e., with respect to the constraints that guide it), a failure to find a ranking in OT can equally result from a small idiosyncrasy (e.g., one letter in which one stroke is produced in the opposite direction than expected) or from a much bigger departure from systematic conformity to constraints. While OT does allow us to investigate cases of failure to model to some degree, it is much more straightforward to do so with HG, as HG allows transparent locating of failures (i.e., which letters/candidates are produced in an inconsistent way compared to the rest of the data), and can provide a degree of fit to the data (i.e., how many letters/targets are classified correctly). In addition to using HG as a way to investigate failures to model with ranked constraints, we also use HG as a benchmark for the performance of OT, and ask whether a weighted sum of constraint violations serves as a better model for human stroke-writing performance than ranked constraints.

Where do constraints come into play?

The framework in which we are working assumes that violable rules, or constraints, govern the way handwritten letters are produced. However, we have thus far not specified where in the process of generating written letters those constraints come into play. In accordance with previous work, we assume that stroke patterns are activated subsequently to the selection of letters and the specification of the exact allographs to be produced (e.g., Palmis, Danna, Velay, & Longcamp, 2017). Most theories of handwriting assume that stored graphic motor plans are activated when a person writes a letter (e.g., Palmis et al., 2017; Rapp & Caramazza, 1997; Van Galen, 1991). Moreover, it is widely assumed that effector-independent graphic motor plans and effector-specific motor programs are distinct, and that the stored, more abstract, effector-independent motor plans are used to compute the effector-specific motor program once the effector is chosen (e.g., Ellis, 1988; Margolin, 1984; Rapp & Caramazza, 1997). We discuss further our work’s contribution to understanding the nature of effector-independent motor plans in the General Discussion.

For this dissertation, we adopt as a working hypothesis that constraints are phrased in effector-independent terms, and come into play as people learn the effector-independent graphic motor plans required to produce character shapes. The learned motor plans form and stabilize not only during initial writing acquisition, when people are taught the prescribed stroke patterns (i.e., in kindergarten or in first grade), but also as they develop their unique handwriting style throughout the first years of writing. Some evidence suggests that the average velocity for writing letters increases with development, making the total writing time shorter (e.g., Palmis et al., 2017; Zesiger, Mounoud, & Hauert, 1993). The increased velocity might be indicative of the establishment of learned motor plans, supporting the idea that writing relies on the retrieval of such stored plans. By proxy, this also supports our hypothesis that constraints come into play as the motor plans become set.

While it seems most likely that the constraints are applied during learning, we do recognize an alternative hypothesis that constraints are activated online every time one produces a letter, dictating in real-time which stroke pattern ought to be used for that letter, and facilitating the generation of a new (rather than retrieving a learned) effector-independent motor plan. This second hypothesis entails perhaps more effort in applying the constraints and their ranking for each new production. We believe it is less likely to be true, but acknowledge it is conceivable. We do note that this hypothesis does not necessarily contradict the option of stored representations. Instead, it implies a different kind of information is stored – rather than complete motor plans, what is stored is the ranking of constraints. Those are then applied when the character to produce is revealed.

Other writing systems

In this dissertation we focus not only on English writing, but on writing in Hebrew as well. Perhaps the most distinct characteristic differentiating Hebrew from English is that Hebrew is read and written from right-to-left. The Hebrew alphabet consists of 22 letters, five of which take a different form when they appear at the end of a word. Hebrew also differs from many other languages in that it has two sets of character shapes which, unlike English upper and lowercase, are not used together (see Table 2). One set is a set of print, or block letters, which we will call "Hebrew print", and the other is a set of handwritten letters. The latter set of characters is called "handwriting" in Hebrew, and sometimes referred to as "cursive Hebrew", even though those letters never connect to one another. For simplicity, and because all the letters in this study are handwritten, we will call this set of characters “Hebrew script”.

Hebrew writing probably developed around the second millennium BCE, and a version similar to the modern Hebrew print form can be found as early as 700 BCE (Sáenz-Badillos, 1993). Hebrew print is used today in formal writing (including the Bible), in typing, in books and newspapers, and online. It is almost exclusively the form people see when reading Hebrew. It is taught in kindergarten or first grade in Israel, where children learn to identify and read the letters. Young children are also taught in school how to write Hebrew print, although by the second grade they will have moved to writing predominantly Hebrew script. Hebrew script is used for informal handwriting (Hebrew script computer fonts exist, but are rarely used). Adult Hebrew speakers in Israel can read both print and script, although they are almost certainly exposed to Hebrew print a lot more than they are to Hebrew script. When typing on a computer, adults usually use the print form, and when handwriting they almost exclusively use script.

LETTER NAME	HEBREW PRINT		HEBREW SCRIPT	
Aleph	א		א	
Beit	ב		ב	
Gimmel	ג		ג	
Dalet	ד		ד	
He	ה		ה	
Vav	ו		ו	
Zain	ז		ז	
Het	ח		ח	
Tet	ט		ט	
Yod	י		י	
Kaf	כ	ך	כ	ך
Lamed	ל		ל	
Mem	מ	ם	מ	ם
Nun	נ	ן	נ	ן
Samach	ס		ס	
Ain	ע		ע	
Pe	פ	ף	פ	ף
Tsadi	צ	ץ	צ	ץ
Qof	ק		ק	
Reish	ר		ר	
Shin	ש		ש	
Tav	ת		ת	

Table 2. The Hebrew letter names, their print form and their script form. Where a letter has a different word-final form, it is displayed on the right.

Summary

In this dissertation we use a novel application of OT and HG to model the stroke patterns in handwritten letters, and to address some of the open questions regarding handwriting. We developed a set of constraints, greatly expanding upon a limited set of rules previously suggested by researchers (e.g., Goodnow and Levine, 1973; Van Sommers, 1984). Our constraints are grounded in fundamental principles of handwriting (principles such as “make your strokes in a way that is well-controlled”, “write efficiently”, “follow the overall direction of reading and writing”, etc.), and are as general and broadly applicable as possible. Using this set of constraints, and applying first the principles of OT, and later of HG, we attempt to model stroke patterns in handwriting. That is, we ask whether a set of violable constraints can account for the way letters are written. We further ask what insights can be gained from this modeling about the cognitive mechanisms underlying handwriting. We also try to gain insight into the representation of constraints, and to highlight opaque relationships between constraints within the same set of characters, and across character sets, languages, writing styles, and effectors.

We first look at the prescribed way of writing, as taught in school, for both English and Hebrew (Chapter 1), and try to model it using OT. Hebrew was chosen as it is different from English in both the shape of letters and the direction of reading and writing, allowing us to examine the role of overall word-writing direction (i.e., the direction of transition from letter to letter) in the way people write and its relation to other constraints or general principles. The first product of our work is showing the extent to which OT can account for the prescribed way of writing letters. We discuss the differences between Hebrew print, Hebrew script, and English prescribed writing, using the differences in constraint rankings between the three character-systems to highlight those differences.

In Chapter 2, we show the extent to which our approach is useful in modeling actual writing data, collected from native English speakers and native Hebrew speakers. In this chapter we further use the framework of OT to evaluate and highlight the differences between the writing of different participants, between prescribed and actual writing, and between writing in different languages. In Chapter 3 we expand upon the previous result by using OT to analyze another set of participants, comprised of right and left-handed writers. Handedness affects the direction in which movement is best-controlled, conflicting, in left-handed individuals writing in English, with the overall writing direction (i.e. for left handed individuals motor-control is better for right-to-left movements, whereas the overall writing direction in English is left-to-right). We use this analysis (in conjunction with right-handed Hebrew speakers' data) to illustrate the resolution of conflicting principles, and to show that our approach is not dependent upon factors such as handedness or overall reading and writing direction.

Finally, in Chapter 4 we use HG to model both the prescribed writing of English and Hebrew, as well as the participant data described in Chapter 3. The purpose of the HG analysis is threefold: 1) to determine whether stratified strict domination is necessary to model the writing-strokes data, or whether small differences among weights are sufficient; 2) to determine whether we can use HG to model the writing in cases where OT modeling fails; and 3) to gain a more straight-forward understanding of cases in which we fail to find a model that accounts for the data. We also use HG to examine the amount of work each of our constraints does, and whether a model can be found using fewer constraints than are needed for OT.

In summary, we ask whether stroke patterns in writing can be modeled with violable constraints that are either ranked in strict domination (OT) or weighted (HG). Given an inventory of characters (e.g., uppercase Roman letters) and an observed stroke pattern for each, we ask whether we can formulate a set of constraints that chooses, for each character, the observed stroke pattern over all other possible patterns. We ask to what extent is our approach useful: Can we model both prescribed and actual writing? In different languages? With different hands? And what can we learn from our success or failure about the cognitive mechanisms underlying handwriting?

Chapter 1: Prescribed writing

In this Chapter, we describe our application of OT to model prescribed handwriting in English and in Hebrew. We modeled the writing of letters according to the way people are taught to write them, typically in kindergarten or first grade. The prescribed way of writing is not subject to some of the "noise" that accompanies actual participants' writing (e.g., due to individual variability, environmental and educational effects, etc.), and so it allowed an initial investigation of the usefulness of the approach to the study of handwriting in a relatively simple setting. We use this chapter to describe our implementation of OT, and make clear some of our assumptions and decisions regarding the definitions of character shapes, targets, and constraints.

Methods

Our modeling using OT included several components. First, we defined the shapes of letters to be produced. Then we systematically generated all the candidate stroke patterns for each shape in each alphabet. From this set of candidates, we chose one designated stroke pattern as the target – the stroke pattern that is taught when children learn to write this letter. If the modeling is to succeed, the target stroke pattern, and only the target stroke pattern, should be chosen as the optimal candidate for each letter. Finally, we defined the set of violable constraints that were used to evaluate the data, and looked for a consistent ranking of them that would choose all and only the targets.

Candidates

For each character system (Roman letters, Hebrew print, Hebrew script), we first defined the set of the basic shapes of the characters that belong to it (e.g., 'A', 'a', 'B', etc.). The set of Roman characters consisted of the 26 letters of the Roman alphabet, in both uppercase and lowercase (for a total of 52 characters). Since Hebrew print and script never appear together in the same text, we defined two separate character sets for Hebrew: One for Hebrew print and one for Hebrew script. Each set consisted of the 27 Hebrew character shapes (22 letters + 5 additional allographs corresponding to the 5 letters that have a different word-final form).

For both Hebrew and Roman characters, several letters have multiple common allographs. For example, the letter a can be written as either 'a' or 'ɑ', and the letter Z can be crossed in the middle

(Ƶ) or not (Z). We added those character shapes to the respective sets of characters, for a total of 67 character-shapes for the Roman alphabet, 33 character-shapes for Hebrew print, and 30 character-shapes for Hebrew script. This list of character shapes was defined based on the prescribed, or "ideal", letter shapes, assuming that even if there are variations in the output shape, the underlying representation remains relatively constant for a given letter.

Letter shapes were determined within the framework often used in typography, with a baseline corresponding to the line on which you write, and an x-line corresponding to the height of regular lowercase letters (Figure 5). Some letters extend below the baseline into the descender area (e.g., the letter 'p'), or above the x-line (e.g., 'b', and all uppercase letters in English), while others are enclosed between those two lines (e.g., 'c').

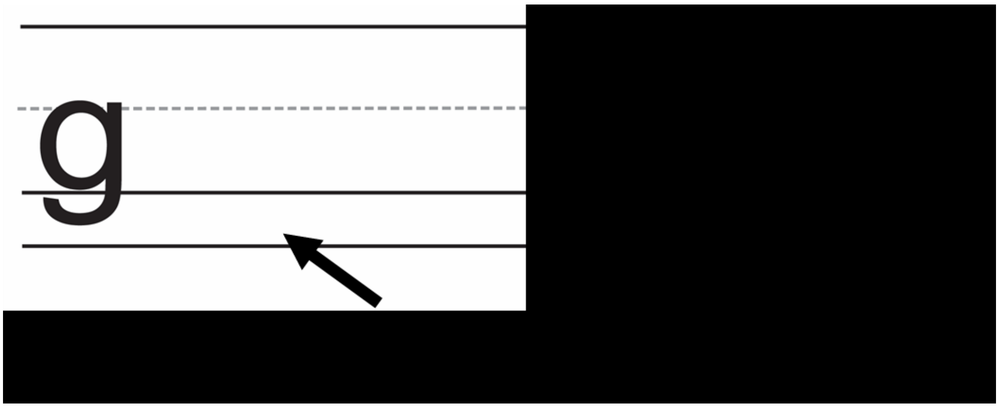


Figure 5. Example of the framework in which letter shapes were defined. Note that the letter 'g' (in the example) extends below the baseline into the descender area.

For each letter-shape, we defined the basic strokes comprising it. A "stroke" was defined as a straight or curved line, or a dot. The ends of a stroke were where the line or curve ended, or where there was a reversal of curvature direction. For example, the letter 'T' consists of one straight horizontal and one straight vertical stroke; the letter 'O' consists of one curved stroke; and the letter B consists of one straight vertical stroke and two curved strokes (Figure 6). Roman uppercase characters had an average of 2.52 strokes per character (SD 0.9), Roman lowercase characters had an average of 2.15 strokes (SD 0.7), Hebrew print characters had an average of 3.0 strokes (SD 0.9), and Hebrew script characters had an average of 1.74 strokes per character (SD 0.6). Character definition always included only the location of the two ends of each stroke, but no information about the direction of production of each stroke, or the order in which strokes are produced.

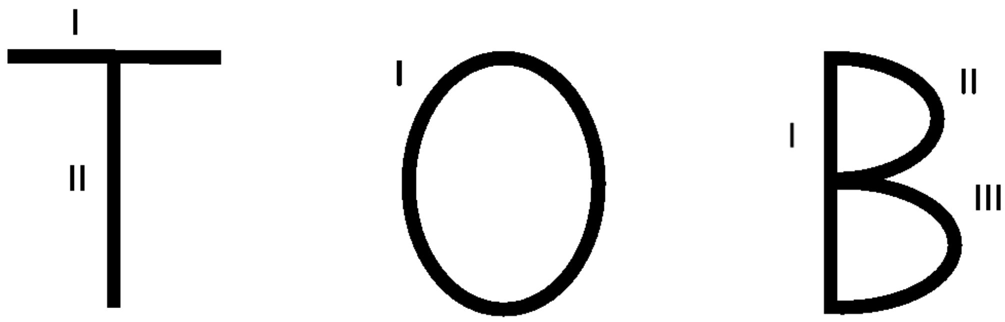


Figure 6. Example character shapes and division into strokes. Left: The letter ‘T’ consists of two straight strokes; middle: The letter ‘O’ consists of one curved stroke; right: The letter ‘B’ consists of one straight stroke and two curved strokes. Roman numbering indicates the number of strokes, not order of production.

The strokes that appear in the character shape are named “base strokes”. Pen movements needed to transition from one stroke to the next (where two strokes are not immediately adjacent at one end) are named “transition strokes” (e.g., the dashed line in Figure 7, left). Transition strokes were not part of the character shape definition, but rather were inserted by a candidate generation function based on the order and direction of base strokes for each candidate. For simplicity, we assumed transition strokes take the shortest path in terms of distance travelled (i.e., a straight line) from the end of one stroke to the beginning of the next. Transition strokes were only part of a candidate when dictated by base-stroke order and direction, and we did not specify zero-length transition strokes. For example, in the production of the letter ‘A’ illustrated in Figure 7a, there is no transition stroke between the first and second base-strokes (as the first base-stroke ends where the second one begins), but there is one between the second and third base-strokes.

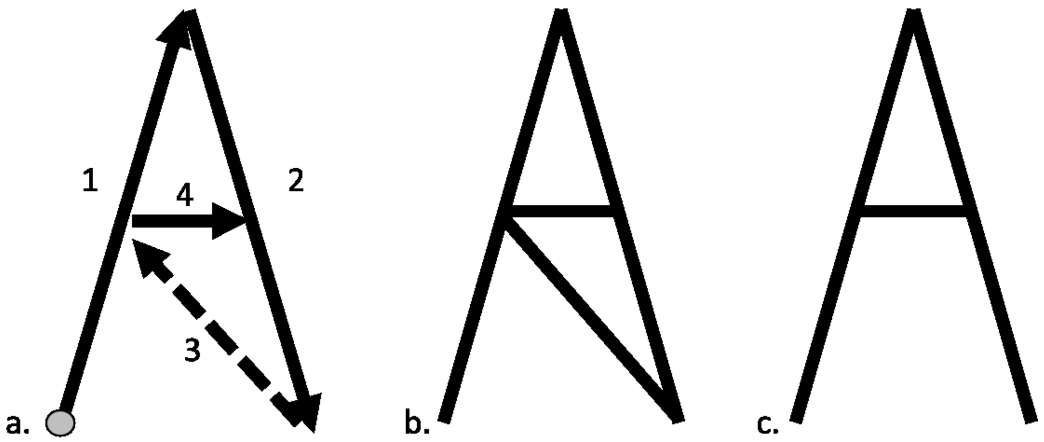


Figure 7. A possible stroke pattern for the letter ‘A’ (a): The gray dot represents the starting point, arrows represent stroke directions, and numbers represent stroke order. The character shape consists of 3 base strokes (labelled 1, 2, and 4), and the order of production in this stroke pattern necessitates a transition stroke (labelled 3). The transition stroke can occur with the pen down, resulting in shape b, or with the pen up, resulting shape c.

For each character, we defined all the possible stroke patterns that could be used to produce its shape. We imposed a few limits on the possible stroke patterns, such as no two consecutive transition strokes, and no starting a stroke other than at one of its ends. The definition of stroke patterns consisted of three things: 1) For each base stroke defined by two end points, in what direction does it proceed (e.g., for the lowercase letter ‘l’, whether the stroke is produced from the top-down or from the bottom-up, and for curves whether they travel clockwise or counter-clockwise). 2) For characters with multiple base strokes, in what order the strokes are produced (e.g., for the letter ‘T’, whether the horizontal line or the vertical line is produced first). And finally, 3) In cases where the order and direction of base-stroke production necessitates a transition between two strokes (as opposed to cases in which one stroke ends at the point where the next begins; e.g., as in Figure 7a), whether the transition stroke occurs with the pen down (resulting in the shape in Figure 7b) or the pen up (Figure 7c). In summary, character shape definitions included only the shape of base-strokes, whereas each candidate stroke pattern, which was computed based on the character shape definition, specified the direction and order of base strokes, as well as the trajectory of transition strokes (the shortest possible line), and whether transition strokes were produced with the pen down or not.

A list of candidates was automatically generated based on the above definitions, using a Matlab script. This corresponds to the GEN function in OT. This process resulted in a list of candidate stroke patterns for each character system, consisting of all the candidates of all the letters in the system. The length of the list of candidates depended on the number of base strokes and transition strokes (e.g., the letter 'O' has two possible candidates – one proceeding clockwise and one proceeding counter-clockwise; the letter 'A' has 176 candidates, depending on the order and direction of all the strokes, and whether transition strokes are pen-up or pen-down). For English upper and lowercase (combined) we had 12,978 candidates, for Hebrew print we had 70,283 candidates, and for Hebrew script we had 694 candidates.

Targets

One of the candidates for each character was designated as the "target" – the stroke pattern that should be chosen as the optimal candidate in the ranking, and have fewer violations of higher-ranked constraints than all other stroke patterns for this character. Targets for this analysis were based on kindergarten textbooks and writing guides from the US and from Israel for English and Hebrew writing, respectively. We only used textbooks intended for native speakers learning to write for the first time (i.e., kindergarten books). Based on those textbooks, which listed the "correct" way to produce a character, we chose for each character the candidate that represents the prescribed way of writing it, and named it the target stroke pattern.

While our target sets fully matched all the textbooks in English and in Hebrew script, there were three characters in Hebrew print ('Pe', 'Shin', and word-final 'Mem') for which there was disagreement between textbooks on the correct way to write them. For two of those letters ('Shin' and word-final ‘Mem’) there was agreement among 4 of the 5 textbooks we surveyed, and so we picked this common stroke pattern as the target. For the letter ‘Pe’ there were three stroke patterns mentioned in the textbooks and guides we surveyed, two of which repeated twice. We chose as the target the stroke pattern that was observed in the two most comprehensive kindergarten textbooks we had.

Constraints

Constraints were intended to be grounded in good control of motion (i.e., hand movements that give better accuracy when using a pen), faithfulness to shape (i.e., the final shape closely resembles the defined ideal letter shape), overall direction of reading and writing (i.e., left-to-right for English, right-to-left for Hebrew), and other general principles suggested by researchers such as Goodnow and Levine (1973), Van Sommers (1984), and others. Based on those principles, we defined a set of constraints, and then employed some trial and error, adjusting to account for language-specific direction, or for certain writing styles or trends we saw in the data. Those language-specific accommodations, as well as the accommodations for different writing styles, occasionally result in seemingly contradicting constraints (e.g., ‘start on the right’ vs. ‘start on the left’). Those contradicting constraints exist together in the set precisely to account for such differences as the difference between writing direction in English and Hebrew, and we did not expect, nor did we observe, any directly contradicting constraints contributing together to a ranking for one set of characters. Eventually we defined a set of 26 constraints which we used to model all of the data throughout this work (Table 3).

CATEGORY	CONSTRAINT
STROKE DIRECTION	no down-to-up base strokes
	no right-to-left base strokes
	no left-to-right base strokes
	no initial down-to-up base strokes
	no down-to-up curved base strokes
	no down-to-up vertical base strokes
STROKE SEQUENCE	no down-to-up sequencing of base strokes (2)
	no right-to-left sequencing of base strokes (2)
	no left-to-right sequencing of base strokes
	minor base strokes after major base strokes
	continue to adjacent base stroke without lifting the pen
	no high-precision intersections
START POSITION	start at leftmost possible start point
	start at rightmost possible start point
	start with stroke containing point closest to upper left corner
	start with stroke containing point closest to upper right corner
	first base stroke starts between baseline & x line (2)
TRANSITION STROKES	no pen lifts
	no transition strokes
	pen up on all transition strokes
	pen up on transition strokes not in base-shape
	transition stroke goes to closer end of next stroke
CURVES	curves continue prior motion direction
	curves counter-clockwise
	curves clockwise
	closed curve start position determines motion direction

Table 3. The Constraints used in this dissertation. A number in parentheses represents our set contains more than one variant of this constraint. See Table 20 in Appendix A for a complete description of each constraint and its variants.

Our constraints were concentrated in a few categories, roughly corresponding to the principles outlined above. The first category pertains to the direction of individual strokes (e.g., whether a vertical stroke is produced from the top-down or bottom-up). Direction of strokes complies with a principle of “good control of motion”, stating that we have better control of the movement when we move the pen from the top-down and (for right-handed individuals) from left-to-right. Good control of motion is important to facilitate legibility, as letter shapes are more legible when the produced strokes are more accurate (or less “sloppy”). Better control of motion probably also contributes to efficiency, as production can be faster when there is better control of the effector, and consistency in the direction of production of strokes (and as a result, consistency in start and end positions), can also contribute to efficiency of the production.

Another category of constraints concerns the sequencing, or the order in which one produces the different strokes (e.g., starting with the vertical, as opposed to the horizontal, stroke in ‘T’), regardless of the direction of each individual stroke. Sequencing-related constraints, as well as the constraints related to the overall starting position (the point or stroke from which you start the entire production of the character) are grounded in both good control of motion and in the principle of efficiency (i.e., writing should be quick and efficient), as writing letters in compliance with the direction of writing a word allows a faster transition between letters, and so does a consistent starting position (since it eliminates the time it takes to compute and reach a variable starting position).

Other constraints pertain specifically to the transitions between strokes. These constraints are grounded in “faithfulness” (another derivative of the legibility principle). For example, we assume that writing (pen-down) on points that are not in the character shape would yield a shape less faithful to the ideal form, thus making it less legible. Similarly, repeating a previously produced stroke (i.e., performing a pen-down transition stroke that merely repeats the trajectory of a previous base-stroke) might create a more “muddied” and less legible shape. Finally, one category of constraints concerns curved strokes specifically. The production of those is not precisely dependent on good control of motion, and may have other cultural and instructional factors contributing to it, as we discuss later.

The way we chose our constraints was motivated by general principles, and still, some decisions had to be made regarding exact definitions of constraints. For example, it seems intuitive that some strokes are more important than others (e.g., the vertical line in lowercase ‘i’ seems to be more important than the dot). Under some accounts, which we adopt, those more important, or "major" strokes should be produced first. However, whether a stroke is a “major” stroke or a “minor” one is not inherent to the strokes, and could be defined in many ways. For example, while most people would agree that the dot in lowercase ‘i’ is a minor stroke, they might differ as to whether the lower horizontal line in uppercase ‘F’ is a minor stroke or a major one. If we include a constraint postulating that major strokes should be produced before minor ones, then the way we define a major stroke will affect the way we compute whether a candidate violates this constraint, and so in a sense, each definition of what a major stroke is, would produce a different variant of this constraint.

All of the constraints we used could theoretically have different variants. In most cases, we managed to find a variant that seemed to work for all of our target sets (in this chapter and the next two). But for three of the constraints in our set we had to use different variants to account for different character sets (e.g., Hebrew print and Roman letters), or for the handwriting of different people (see Chapters 2 and 3). For the constraints ‘first base stroke starts between the baseline and x-line’, and ‘no right-to-left sequencing of strokes’ we had one main variant that was useful almost across the board (for the prescribed writing and 27 or 28 participants, respectively), and a secondary variant that was used by 5 and 2 participants, respectively. For the constraint ‘no down-to-up sequencing of base strokes’ we also had two variants, one used by 7 participants (all English speakers), and the other used by 3 participants (all Hebrew speakers; see Chapter 2). In each of the three cases, we named the more common variant “v1”, and the less-common one “v2”. We acknowledge the possibility that a universally-applicable variant exists for each of the above three constraints, which we simply have not found yet.

Data analysis

For each candidate stroke pattern for each character we generated a profile of constraint violations by tallying the number of violations of each constraint. This resulted in a tableau of constraint violations for each character (with one row for each candidate and one column for each constraint). When pooled together, the resulting matrix consisted of as many rows as there were candidates in the language. For example, for English, there were 12,978 rows, corresponding to the same number of candidates, and 26 columns, corresponding to our set of constraints. Each cell in the matrix indicated the number of times this particular candidate violated this particular constraint (Figure 8).

	Constraint violations			
Candidate	[redacted]			
1	**0**	**2**	**0**	**…**
[redacted]				
3	**1**	**2**	**2**	**…**
[redacted]				
12,978	**1**	**3**	**0**	**…**

Figure 8. Example constraint violation matrix. Each row represents one candidate stroke pattern, and each column represents one constraint. Each cell notes the number of times a constraint is violated in a given candidate stroke pattern. Note that the matrix includes all the candidates of all the characters in a system, not just a single character.

To determine whether there is a successful ranking of constraints that would yield as the optimal outputs only those candidates we designated as targets, we implemented the recursive constraint demotion algorithm (Tesar & Smolensky, 1998), which is based on the principle of strict domination. This algorithm guarantees that required domination relationships are respected, by “demoting” constraints into the highest rank possible that will not "choose" a non-target over a target. It ensures a successful ranking will be found if such a ranking exists, and if there is no ranking of the input constraints that would choose all and only target candidates, the algorithm fails.

We implemented the algorithm in Matlab and ran it separately on each set of characters (Hebrew print, Hebrew script, and Roman letters). For each set, the algorithm ran on all characters in the set, first ranking all constraints as equal in the highest stratum (i.e., in the highest rank), and then demoting constraints that chose non-target candidates over the target candidate for each letter. If the target candidate had more violations of a constraint than a non-target candidate did, this constraint was demoted in the ranking until it was ranked below another constraint that chose the target over the non-target. If no such constraint was in the set, or if there was a conflict in domination relations between two letters, the ranking algorithm would fail, and the conclusion would be that there does not exist a ranking of these constraints that accounts for all and only target stroke patterns. After running the constraint demotion algorithm, if the resulting ranking had each target stroke pattern "winning" over all its non-target alternatives (i.e., all target sequences for each letter were more Harmonic than all non-target sequences for that letter), then the algorithm is said to have succeeded in finding a consistent ranking.

The constraint demotion algorithm is not guaranteed to produce a unique ranking. Instead, it ranks constraints by strata, with a ranking on a lower stratum meaning that at least some constraint was required to be ranked above the constraint from the lower stratum. Placing constraints in the same stratum indicates that the data do not require those constraints to be ranked differently, not that they are equally important or that they necessarily need the same ranking. As a general rule, the rank of constraints does not imply importance or the constraint's usefulness in achieving a successful ranking. A constraint ranked at the highest stratum could be the most important one in a sense, if it is required to dominate all the lower-ranked constraints, but it could just as well not be needed at all, either because it did not apply to any character, or because some other constraint already picks all the winners for which this constraint would be required. Furthermore, two constraints could be ranked in the highest stratum with one of them having to dominate all other constraints and the other having to dominate only one other constraint or a handful of them. A constraint may also never get demoted by the algorithm, thus staying in the highest stratum, despite not being necessary for a successful ranking.

Because of the under-specification of the ranking, it was important to prune out those constraints that were not needed for a ranking at all. Assuming that a successful ranking of constraints was found, we attempted to find the minimal set of constraints that were sufficient for a ranking. To find the minimal set of constraints we implemented an algorithm in Matlab that operated in two stages: We first determined if there were any constraints that were singly necessary to achieve a successful ranking. That is, the algorithm searched for any target candidates that had an equal or greater number of violations than a non-target candidate on all but one constraint. In the second stage, we ran the ranking algorithm to check if there is a consistent ranking that would account for all and only target stroke patterns using only this subset of constraints that were singly necessary. If a ranking was found, this set was deemed the minimal constraint set (i.e., the smallest set of constraints that can account for the particular set of characters and targets). If none was found however, we added one of the remaining constraints from the original list of 26 (one that was not already a singly necessary constraint), and ran the algorithm again. We repeated this process, each time adding one of the remaining constraints, until we exhausted the set. If still no ranking was found, we ran the algorithm again, this time adding to the set of singly necessary constraints not one but two of the remaining constraints, and so forth, until a successful ranking was achieved.

The constraint demotion algorithm determines whether there is at least one ranking that would successfully account for all and only target candidates given a set of constraints, but it does not reveal which constraints are necessary or sufficient for success, nor does it reveal individual constraints' importance or usefulness for the ranking. After determining the minimal set sufficient for a ranking, we also wanted to find out what necessary domination relations exist between the ranked constraints. To that end, we implemented the Fusional Reduction algorithm (FRed; Brasoveanu & Prince, 2005). This algorithm takes the profile of constraint violations by all candidates in a set, and finds which domination relations are required to achieve a successful ranking. FRed reveals both the necessary and sufficient domination relations between constraints to find a ranking given a data set and a set of constraints. In conjunction with our minimal constraint set script, we were able to find both the smallest set of constraints sufficient for a ranking and the necessary domination relations between them.

Summary of data analysis

After defining all the candidates and target stroke patterns for each character set (Roman letters, Hebrew print letters, Hebrew script letters), we ran the constraint demotion algorithm on each set separately. In each run, we included the 26 constraints (detailed above), and all the allographs for each character set. If a successful ranking was found, we ran the additional scripts to determine the minimal set of constraints that is sufficient to account for the data, and the necessary domination relations between constraints.

Results

The ranking algorithm was run separately on the set of Roman characters (upper and lowercase combined), on the set of Hebrew print characters, and on the set of Hebrew script characters. In each of the three cases we managed to find a ranking of the 26 constraints that would account for all and only target stroke patterns. After running the minimal-constraints algorithm, we found that Roman characters could be accounted for using a minimal set of 12 constraints (Table 4, top). Hebrew print could be accounted for using 11 constraints (6 of which were shared with the Roman set; Table 4, middle), and Hebrew script could be accounted for using 7 constraints (3 shared with Hebrew print and 3 shared with Roman; Table 4, bottom).

Table 4. Minimal constraint set and ranking (by stratum) that account for all and only target stroke patterns for prescribed writing of Roman upper and lowercase (top), Hebrew print (middle) and Hebrew script (bottom). If a certain constraint had multiple versions in our set, the version used is noted in parentheses. Relative rank within stratum is not determined in the above table.

CONSTRAINT	RANK
ROMAN UPPER AND LOWERCASE	
pen up on all transition strokes	1
no initial down-to-up base strokes	1
no down-to-up curved base strokes	1
minor base strokes after major base strokes	1
closed curve start position determines motion direction	1
start at leftmost possible start point	2
no down-to-up sequencing of base strokes (v2)	3
no right-to-left sequencing of base strokes (v1)	4
start with stroke containing point closest to upper left corner	5
no pen lifts	5
no down-to-up base strokes	6
no right-to-left base strokes	7
HEBREW PRINT	
pen up on all transition strokes	1
curves clockwise	1
continue to adjacent base stroke without lifting the pen	1
no down-to-up vertical base strokes	2
no pen lifts	3

no left-to-right sequencing of base strokes	4
no down-to-up base strokes	5
start at leftmost possible start point	6
no right-to-left base strokes	6
minor base strokes after major base strokes	7
start with stroke containing point closest to upper right corner	8
HEBREW SCRIPT	
pen up on all transition strokes	1
no high-precision intersections	1
first base stroke starts between baseline & x line (v1)	1
closed curve start position determines motion direction	1
no down-to-up base strokes	2
start with stroke containing point closest to upper right corner	3
start at rightmost possible start point	4

FRed analyses were run on each of the minimal sets to determine which domination relations were necessary for a successful ranking. This allowed us to make judgements about the importance and participation of individual constraints, including differentiating constraints that were ranked in the same strata. For example, in the ranking for Roman prescribed, the constraint ‘minor base strokes after major base strokes’ was ranked in the highest stratum along with ‘closed curve start position determines motion direction’, but the former was required to dominate six other constraints whereas the latter was not required to dominate any other constraints. Thus, two constraints, which on the surface had the same ranking, were in fact very different in terms of their participation in the modeling. Interestingly, the constraint ‘minor base strokes after major base strokes’, which was ranked in the highest stratum for Roman letters, and had to dominate six other constraints, was ranked in the seventh (and second-lowest) stratum for Hebrew print, where it had to be dominated by eight constraints, and only dominate one. We discuss further individual constraints' contribution to the ranking in Chapter 4.

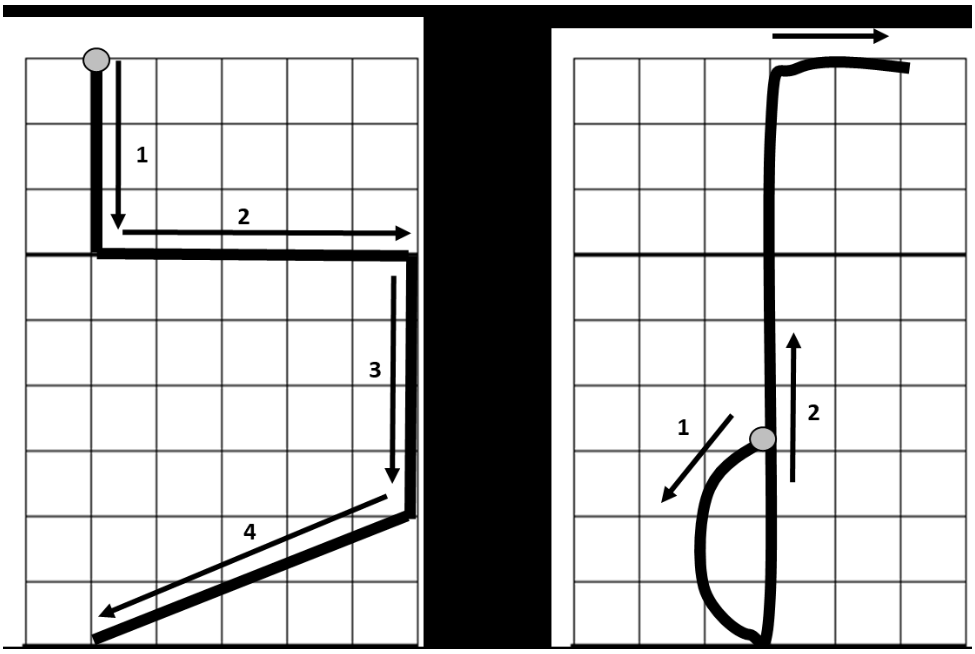
Using FRed, we were also able to directly compare domination relations among the three character-sets examined in this analysis (Roman, Hebrew print, and Hebrew script). Some similarities were evident among the three. For example, they all needed 'pen up on all transition strokes' to be part of their minimal set to achieve a successful modeling, and all three sets required that this constraint dominate 'no pen lifts'. Other domination relations stood directly opposed when comparing two languages, and those differences can be traced back to characteristic features of the language. One of these contrasts concerns high-precision intersections: places in which two strokes meet, and the character shape dictates that the second stroke pass through (or arrive at) a particular point in the first stroke. For example, in the uppercase letter T, if the horizontal stroke is produced second, it must pass through the end-point of the vertical stroke (a high-precision intersection). In a lowercase t, on the other hand, the horizontal stroke can pass through any point within a section of the vertical stroke, making this a low-precision intersection.

A constraint dictating 'no high-precision intersections' would facilitate faster production. Therefore, in scripts or in settings that prioritize writing quickly or efficiently over writing accurately or legibly, this constraint might have to dominate constraints grounded in good control of motion, such as 'no down-to-up strokes.' And indeed, this is the case in Hebrew script. Hebrew script, unlike Hebrew print, was developed with the intention of being handwritten, and it therefore puts a much greater emphasis on writing quickly and efficiently than print does. Since high-precision intersections are slower and less efficient to produce, Hebrew script prioritized 'no high-precision intersections' over 'no down-to-up base strokes', requiring the former to dominate the latter. Conversely, Hebrew print and English prescribed writing, both optimized for accuracy, needed the opposite domination relation, dictating that strokes be produced in the direction that allows better control over the movement.

Another example of the trade-off between accuracy and efficiency can be seen with the constraints 'no down-to-up base strokes' and 'start between the baseline and the x-line'. The ranking of the constraints is such that 'no down-to-up strokes' has to dominate 'start between the baseline and x-line' in Hebrew print, but the reverse domination relation is necessary to model Hebrew script. This fits with the usage of each writing system: Hebrew print is formal, and so values accuracy and legibility over speed (producing strokes from top to bottom means better control of motion, and thus better accuracy in stroke production, leading to improved legibility), whereas Hebrew script is informal, and values efficiency and speed (starting between the baseline and x-line means more consistency in starting position, allowing a quicker transition between letters).

This conflict in necessary domination relations can be demonstrated with the pair of characters below (Figure 9). When produced according to the prescribed stroke pattern, the character on the left, Lamed in Hebrew print, starts at the top (above the x-line), thus violating the constraint 'start between the baseline and x-line'. On the other hand, in its prescribed form, the character on the right, Lamed in Hebrew script, is produced starting with the loop in the middle, which requires it to then violate the constraint 'no down-to-up strokes' when producing the next stroke. The fact that these two characters share the same letter identity throws the contrast between the two character-sets into sharp relief.

Figure 9. The letter Lamed in Hebrew print (left) and in Hebrew script (right). Hebrew print prefers 'no down-to-up strokes' over 'start between the baseline and x-line', whereas Hebrew script requires the opposite domination relation. Arrows represent stroke direction, the gray dot represents starting position, and the numbers represent stroke order. The bottom border of the grid is the baseline and the thick black grid-line is the x-line.



Discussion

In this chapter, we looked at the prescribed way of writing letter-strokes in English (Roman upper and lowercase), in Hebrew print, and in Hebrew script. We asked whether we can account for the prescribed stroke patterns using ranked violable constraints in all three sets of characters, what the similarities and differences between the sets are, and what can be learned from those similarities and differences. Our first goal in this analysis was to find out whether there is a consistent ranking of constraints that can account for the stroke patterns of all and only target letters for Roman characters (upper and lowercase), for Hebrew print, and for Hebrew script. For each of those sets we have found a ranking of constraints that chooses all and only target stroke patters. The fact that the prescribed way of writing both English and Hebrew can be modeled using ranked violable constraints suggests that there is systematicity in the way letters are produced, and that this systematicity can be captured within the framework of OT.

Simply modeling our data was not our only goal. We also wanted to use OT and the ranking of these constraints to reveal complex and opaque relationships between principles that come into play when writing, as well as how they manifest with different constraints, in different languages, or even different characters within a set. These more specific insights could only be drawn once we have established that the framework is viable. After finding a ranking the chooses all and only target stroke patterns over non-target patterns, we ran the Fusional Reduction algorithm (FRed) on the minimal set of constraints for each language, and the combined set of constraints for every pair of languages. Using FRed, we were able to find necessary domination relations, and to differentiate which constraints are, in a sense, more important within a ranking (and even within the same stratum), and across the languages.

We discovered, in modeling Hebrew print and script, that there are some inherent conflicts that prevent these two scripts from being modeled with the same ranking of constraints. This might seem unsurprising, as the two sets are never used together in the same text (unlike English upper and lowercase), but until we examined the necessary domination relations using FRed, we did not know the exact nature of these conflicts, nor could we say for sure that these two character-sets cannot be modeled together. Once we analyzed the results of FRed, we found that six pairs of constraints stood in direct contradiction in terms of their required domination relations. These contradictions mean that there could not possibly be a ranking of those constraints within our framework that would account for target stroke patterns for both Hebrew print and Hebrew script modeled together.

Clashing principles: Writing direction vs. control of motion

One of the questions that has been raised regarding Hebrew writing is how to reconcile two principles governing writing direction: Good control of motion (which in right-handed individuals is better for left-to-right horizontal movements of the pen), and the direction of progression from character to character (which complies with the reading direction, right-to-left in Hebrew). Consider for example the letter Het (ח; /xet/) in Hebrew print. One possibility is that the production of the individual letter entirely follows the overall direction of reading and writing in Hebrew (right-to-left), and thus the production of the letter Het would progress from right-to-left, as in Figure 10a (left). However, for right handed individuals, good control of motion dictates that a stroke with a horizontal component be produced from left-to-right. Therefore, some researchers (e.g., Van Sommers, 1984) have suggested that each individual Hebrew letter is produced entirely from left-to-right (as in Figure 10b), despite the direction of progression between letters being from right-to-left.

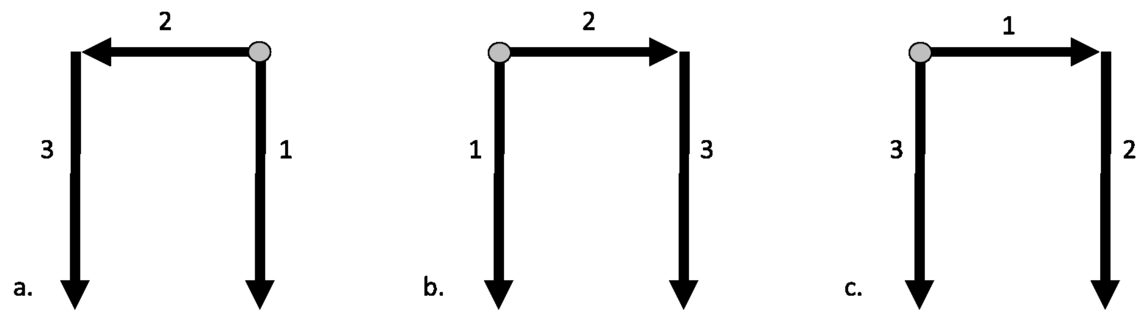


Figure 10. Three possible ways to produce the letter Het in Hebrew print: A stroke pattern that favors both R-L horizontal strokes and R-L sequencing of strokes (a); A stroke pattern that favors both L-R horizontal strokes and L-R sequencing of strokes (b); And a Stroke pattern that favors L-R horizontal strokes, but R-L sequencing of strokes (c). Arrows represent stroke direction, the gray dot represents starting position, and the numbers represent stroke order.

In English there is no conflict, and so a production that follows the direction of reading and writing (left-to-right) also satisfies the requirement for good control of motion on strokes with a horizontal component (for right-handed individuals). Accordingly, for the Roman characters we see that the constraints needed for a successful ranking are ‘no R-L base strokes’ (applied to strokes with a horizontal component) and ‘no R-L sequencing of strokes’. In Hebrew, the reality revealed by our analysis is more complex. The target stroke pattern for prescribed writing of the letter Het has the horizontal stroke produced from left-to-right, and the two vertical strokes sequenced such that the one on the right is produced first and the one on the left produced second (Figure 10c). We found that in order to model Hebrew print writing successfully, both ‘no R-L base strokes’ and ‘no L-R sequencing of strokes’ needed to be ranked.

As emerges from our analysis, while there is a constraint imposing that horizontal strokes in Hebrew should generally be produced from left-to-right, in accordance with good control of motion, the sequencing, or the order of the strokes proceeds from right-to-left, bringing the last stroke closest to the next letter to be produced. This relationship was not transparent; there are always violations of constraints, and indeed some strokes in the prescribed writing of Hebrew are produced from right-to-left, and occasionally there is left-to-right sequencing of strokes (e.g., when the production of a particular letter dictates compliance with some higher-ranked constraint). However, by using the framework of OT, we were able to reveal that both of these constraints are in fact necessary for the modeling, and that the constraint governing the sequence, or the order of strokes must dominate the constraint governing the direction of individual strokes. This domination relation is perhaps the catalyst behind the erroneous notion that Hebrew letters are produced entirely from right-to-left.

This example illustrates how the writing system comes to grips with conflicting principles. When two principles, in this case writing in a direction that conforms with the overall direction of reading and writing in a language, and writing in a way that allows best control over the motion, both contribute to the writing of a language, the conflict between them needs to be resolved. One way to deal with the conflict is to "explain it away," by abandoning one of the principles altogether. However, if this was the case, we would not observe the abandoned principle in the rules governing the language at all. Researchers in the past seemed to dodge the issue of clashing principles, by assuming that both principles apply, most of the time. Using ranked violable constraints, we can show how both principles indeed apply most of the time, but also how the hierarchy between them allows a resolution of the conflict when one (or both) has to be violated.

Chapter 2: Participant writing – Hebrew and English

In this chapter we apply our OT framework to the writing of native English speakers and native Hebrew speakers. We attempt to model their handwriting using the same framework and the same constraints described earlier. We ask whether the approach is useful not just for the modeling of the relatively structured prescribed writing, but also for the jungle that is actual human performance. We further use the modeling to try to explain the underlying causes of variability between participants.

Methods

Participants

Twenty-four native English speakers (9 females, 15 males), and 24 native Hebrew speakers (11 females, 13 males) with no known language or neurological impairments participated in this study. Participants gave written informed consent in accordance with the Johns Hopkins Institutional Review Board and were offered either course credit or $10 compensation for their time. English speakers had a mean age of 23;0 years (SD 11;11, median 19;11), and Hebrew speakers had a mean age of 35;7 years (SD 14;0, median 30;4). Age was significantly different between the two groups, $t(46) = 3.287$, $p = 0.002$, however, since we are not directly comparing the groups on any measure, we do not believe this has an impact on the results. English speakers had completed on average 14.9 years of education (median 14), and Hebrew speakers 16.5 years of education (median 16). Four of the English speakers and one of the Hebrew speakers were left-handed. Two of the participants of the Hebrew group were native speakers of English as well as Hebrew. All of the participants in each group had their respective language as their first language of instruction (see summary in Table 5).

	ENGLISH SPEAKERS	HEBREW SPEAKERS
N	24	24
FEMALE/MALE	9/15	11/13
MEAN AGE (SD)	23;0 (11;11)	35;7 (14;0)
MEDIAN AGE	19;11	30;4
MEAN EDUCATION, YEARS (SD)	14.9 (2.3)	16.5 (3.3)
MEDIAN EDUCATION, YEARS	14	16
RIGHT HANDED/LEFT HANDED	20/4	23/1

Table 5. Summary of demographic details for the Hebrew and English participant groups.

Stimuli and procedure

For the English speaker group, stimuli consisted of 85 mono-morphemic English words (average length 6.2 letters). The list was designed to include each letter in initial and in non-initial position at least twice. Each letter appeared on the list a minimum of 5 times (Z) and a maximum of 53 times (E). The list of 85 words was divided into two subsets (of 43 and 42 words), and each subset was dictated to participants twice. For each subset of words, participants were asked to repeat each word, and then write it in either all uppercase or all lowercase letters, depending on the subset (the order of words within each list as well as the order of upper or lowercase was counterbalanced across participants). After each subset of English words, participants were given a copying task of Hebrew letters. Each letter was presented on a single sheet of landscape-oriented A4 paper in 54-point font (the font used was Arial for Hebrew print and Guttman Yad-Brush for Hebrew script). The list consisted of the 27 Hebrew letter forms (22 letters + 5 word-final forms), once in Hebrew print and once in Hebrew script. Each list (Hebrew print and Hebrew script letters) was divided into two subsets, and the order of the subsets was counter-balanced across participants. Participants were told the stimuli were “letters in another language,” and asked to copy them onto their page.

For the Hebrew speaker group, stimuli consisted of 32 mono-morphemic Hebrew words (average length 3.6 letters), as well as 26 mono-morphemic English words (average length 4.3 letters). The list of Hebrew words was designed to include each letter in initial, medial, and final position in the word at least once, and letters that have a different word-final form appeared at least twice in that position. English words had each letter in initial and non-initial position at least once. Each list of words was dictated to participants twice (in different orders, counterbalanced across participants). Participants wrote the Hebrew words once in print and once in script (depending on the list), and the English words once in all uppercase and once in all lowercase (depending on the list). The order of the lists was counter-balanced across participants. In addition, 14 of the 24 Hebrew speakers were given the same copy task of Hebrew letters as the English participants (see description above). They were instructed to copy the letters in either print or script as presented. The letters were copied rather than written to dictation to allow a future comparison to the copying of the same letter shapes by the English speaker group.

For each list of words (English words for the English speaker group, English and Hebrew words for the Hebrew speaker group), the words were dictated by the experimenter, and participants were asked to repeat each word out loud before writing it down. Participants wrote using a specialized inking pen, on regular paper that was placed on a Wacom Intuos Pro 3 graphics tablet. Participants were instructed to write within the surface of the tablet, and to not connect letters (letters in Hebrew do not connect, but participants were nevertheless reminded of this instruction before each list of words, regardless of language or case). They were also encouraged to use as much space as they needed, and not to write in too small a font. Participants were told that if they made a mistake in the spelling of a particular word, or produced a different shape than they intended (e.g., by starting a stroke too far to the top or bottom, or by accidentally omitting a stroke), they could ask to write it again. This happened only rarely. Participants were debriefed about the purpose of the experiment after their participation, and filled questionnaires for handedness and relevant biography.

Candidates and constraints

In the current analysis, we focused on the writing of only English lowercase letters (for the English group) and Hebrew script letters (for the Hebrew group), as these are produced more often, and are more likely to be stable within participant. Candidate stroke patterns were defined the same way as in Chapter 1. Some letters have different acceptable allographs (e.g., ‘z’ can be written with or without a horizontal crossing line in the middle, and ‘a’ can be written using the one-story or two-story allograph). We have therefore generated candidate stroke patterns for the entire set of allographs we observed, for a total of 41 allographs for English lowercase and 31 allographs for Hebrew script. The constraint set consisted of the same constraints described above in Chapter 1.

Targets and data analysis

All analyses were carried out using Matlab. Participant writing data was collected and extracted using either Ductus (Guinet & Kandel, 2010), or Matlab scripts written especially for this purpose. For each participant, we determined the stroke pattern that was used for each letter. This stroke pattern was considered the target for this character for this participant. Occasionally, a participant produced a letter in two different ways on different trials. This happened on 23 characters in total across the 24 English speakers, and on 6 characters in total across the 24 Hebrew speakers. In those cases, we marked both observed candidates as targets, and a successful ranking would mean both targets won over all alternative non-target candidates for that letter (but we did not specify that both targets have to tie, or that one has to be more Harmonic than the other). In cases where participants produced one allograph instead of another, we excluded the allograph that was never observed. The constraint demotion algorithm was run on the set of targets for each participant for each character set. When the ranking was successful, we also ran the Matlab program determining the minimal set of constraints for that participant (as described in Chapter 1).

Results

We attempted to find a ranking of constraints that would choose all and only target stroke patterns for English participants writing in lowercase and for Hebrew participants writing in script. We managed to find such a ranking for eight of the 24 English speakers, and for 18 of the 24 Hebrew speakers, including the only Hebrew-speaking left-handed participant in this experiment. Finding a ranking of constraints that fully accounted for the handwritten stroke-patterns of actual participants was not trivially derived from the success of the prescribed modeling. Every single participant deviated from the prescribed stroke patterns in at least 2 letters (in Hebrew), or 13 letters (in English). The average number of letters with stroke patterns different from the prescribed form was 4.2 letters in Hebrew script (range 2-8), and 15.3 in English lowercase (range 13-22).

Of the participants whose handwriting we could not fully model, with some we came closer than others. For many participants, only one target was incompatible with a ranking that accounted for all other letters. In other words, for those participants, only one character had a candidate ranked higher than the target stroke pattern, and all other targets were ranked higher than their alternatives. It is possible that there is something fundamental that cannot be reconciled in the way those participants write, and that they generally do not adhere to rules of the type we have been discussing (i.e., ranked violable constraints). However, since it was only one letter for each of those participants that was incompatible with the rest, we reasoned that perhaps their writing of some letters is occasionally idiosyncratic; the incompatible letters are written a certain way even though their stroke pattern does not fit the overall scheme (we discuss this point further in Chapter 4 and in the General Discussion). We defined a cutoff of one letter, which, when this letter was excluded, yielded a successful ranking, and named this a “partial success”.

For those participants whose ranking only failed on one letter (i.e., for whom we achieved partial success), we ran the ranking algorithm again excluding only this letter. This allowed us to partially account for the handwriting of eight additional English speakers, including two of the four left-handed participants (for a total of 16; 14 right-handed), and four additional Hebrew speakers (for a total of 22; 21 right handed).

English-speaking participants needed a minimal set of between 8 and 12 constraints (average 9.7, median 10) to yield a successful ranking, and Hebrew-speaking participants needed a minimal set of between 7 and 10 constraints (average 7.9, median 8). We needed a combined set of 19 constraints to account for the writing of all 16 English-speaking participants (Table 6, top), with three constraints being used by every single participant. Two constraints ('first stroke starts between baseline and x-line', and 'no right-to-left sequencing of strokes') were used in a different version by different participants. For Hebrew script, we needed a combined set of 13 constraints to account for the writing of all 22 Hebrew-speaking participants, with four constraints being used by every participant (Table 6, bottom).

Table 6. The constraints used in the modeling of English speakers' writing of lowercase Roman letters (top), and Hebrew speakers' writing of Hebrew script letters (bottom). If a certain constraint had multiple versions, the version used is noted in parentheses. On the right: For how many participants (out of 16 for the English group and 22 for the Hebrew group) was this constraint in the minimal set.

CONSTRAINT	PARTICIPANTS
ENGLISH LOWERCASE	
pen up on transition strokes not in base-shape	16
no pen lifts	16
no right-to-left base strokes	16
no down-to-up base strokes	14
minor base strokes after major base strokes	12
no transition strokes	12
closed curve start position determines motion direction	11

start with stroke containing point closest to upper left corner	10
no right-to-left sequencing of base strokes (v1)	8
curves continue prior motion direction	5
curves counter-clockwise	4
start at rightmost possible start point	3
no down-to-up vertical base strokes	3
no down-to-up sequencing of base strokes (v2)	3
first base stroke starts between baseline & x line (v2)	2
start at leftmost possible start point	2
no right-to-left sequencing of base strokes (v2)	2
no initial down-to-up base strokes	1
first base stroke starts between baseline & x line (v1)	1
HEBREW SCRIPT	
pen up on transition strokes not in base-shape	22
no down-to-up base strokes	22
first base stroke starts between baseline & x line (v2)	22
no high-precision intersections	22
start with stroke containing point closest to upper right corner	21
curves clockwise	19
no pen lifts	18
start at rightmost possible start point	11
start at leftmost possible start point	5
no down-to-up sequencing of base strokes (v1)	3
closed curve start position determines motion direction	2
curves counter-clockwise	2

no down-to-up curved base strokes	1

We did not find a difference between our ability to model the handwriting of participants who only produced one stroke pattern for each character and those who produced some characters in multiple different ways on different trials, as demonstrated with Chi square tests for dependence. Of the English speakers, 14 of the 24 participants produced at least one letter in more than one way across the testing session, and we were able to model the handwriting of 4 of them (χ^2 = 0.59, p = 0.56). Of the Hebrew speakers, 6 of the 24 participants produced at least one letter in more than one way, and we were able to model the handwriting of 5 of them (χ^2 = 0.67, p = 0.59). We also did not find any difference in the number of letters whose stroke patterns deviated from the prescribed writing between the participants whose writing we could model and those we could not (English: $F(1, 22)$ = 1.78, p = 0.20, Hebrew: $F(1, 22)$ = 0.95, p = 0.34).

Explaining variability across participants using constraint rankings

Using ranked, violable constraints we were able to account for the handwriting of different participants, who used different sets of stroke patterns. We found that different constraint rankings gave rise to, and explained, the variability in stroke patterns among participants. For example, we found that a small change in constraint rankings could explain the difference between the handwriting of two participants who produced only one character differently from one another. GEM and LLN, two Hebrew speaking participants, produced every letter the same way in Hebrew script, except for the letter Shin (which looks like a Roman lowercase ‘e’; Figure 11). While GEM produced Shin from the middle-left and going counter-clockwise (as one might produce a lowercase e; Figure 11a), LLN produced it from the bottom-right and going clockwise (as is the prescribed stroke pattern in Hebrew; Figure 11b).

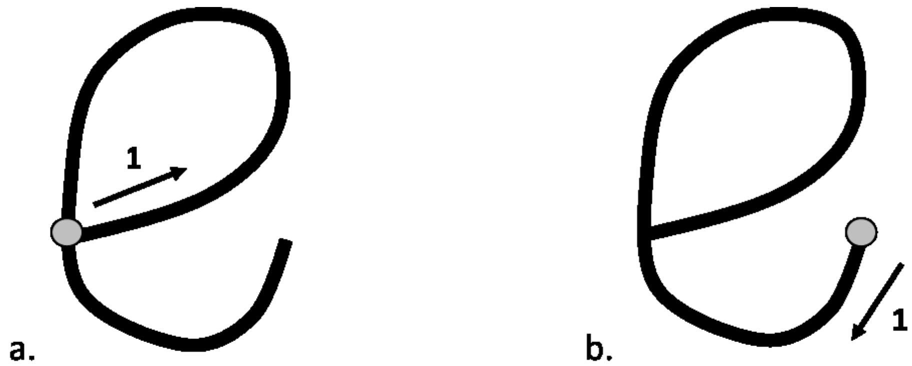


Figure 11. Hebrew script Shin as produced by participant GEM (a) and LLN (b). The gray dot marks the starting position, and the arrows mark the direction of production.

When looking at their ranked constraints, both GEM and LLN used the same minimal set of eight constraints (Table 7). However, GEM's handwriting was modeled with a ranking in five strata and LLN's with six. This in itself is of course not very informative, as the strata alone do not necessarily represent what domination relations are needed. We therefore ran FRed on both participants' tableaus to find the differences in their necessary domination relations. We found that all the domination relations between the two participants were the same, except that GEM needed 'start at leftmost possible start point' to dominate 'curves clockwise', whereas LLN needed the opposite domination relation to account for her writing. Of course this explains the difference in stroke patterns for Shin, as producing the curve clockwise would necessarily entail not starting it on the left, and vice versa.

GEM		**LLN**	
Str.	**Const.**	**Str.**	**Const.**
1	pen up on transition strokes not in base-shape	**1**	pen up on transition strokes not in base-shape
	no high-precision intersections		no high-precision intersections
	first base stroke starts between baseline & x line		first base stroke starts between baseline & x line
2	no D-U base strokes	**2**	no D-U base strokes
3	start at leftmost possible start point	**3**	curves clockwise
4	curves clockwise	**4**	start at leftmost possible start point
	start with stroke containing point closest to upper right corner	**5**	start with stroke containing point closest to upper right corner
5	no pen lifts	**6**	no pen lifts

Table 7. The minimal constraint set used to model the writing of participant GEM (left) and LLN (right). The two participants use the same constraints, ranked in a different order. Str. = stratum; Const. = constraint. See Table 20 in Appendix A for a complete description of each constraint.

The example of GEM and LLN is a simple and clear demonstration of how a small change in domination relations can effect a small change in stroke patterns. Note that the dominating constraint is ranked relatively low (in the 3rd stratum out of 5 for GEM and out of 6 for LLN; with the dominated constraint obviously ranked lower), thus creating a relatively small variation when flipping the domination relations. Following our success with explaining a small difference in stroke patterns, we looked at the difference in domination relations between participants who differed on a larger number of letters.

RYI and AAL, two participants from the English-speaker group, differed in the stroke patterns they produced on five Roman lowercase letters (d, k, p, x, and y), and produced different allographs on two more (u and z). RYI and AAL each needed a minimal set of 10 constraints to account for their writing. Despite the relatively large differences in their stroke patters (different targets in 19% of letters), 9 of the constraints in RYI and AAL's minimal sets were shared, and only one constraint differed between them: RYI used 'no right-to-left sequencing of strokes', whereas AAL used 'start at rightmost possible start point' (Table 8). The odd usage of a starting point closest to the right for an English speaker was required to explain AAL's production of both x and d starting at the top-right. While there are a few obvious similarities between the rankings of RYI and AAL (e.g., they both require 'pen up on strokes not in base shape' to be ranked highest), and despite them using an almost identical set of constraints, the ranking of constraints in itself does not reveal the full picture of the differences between them.

Table 8. Ranking of constraints used to model the writing of participant RYI (left) and AAL (right). Where a constraint has multiple versions, the version is indicated in parentheses. A ranking of 0 indicates the constraint does not get ranked. Where a constraint is not required to be in the minimal set, it is marked with italics and an asterisk. The two participants share nine constraints, and their constraints are ranked in a different order. Str. = stratum; Const. = constraint. See Table 20 in Appendix A for a complete description of each constraint and its variants.

RYI		AAL	
Str.	**Const.**	**Str.**	**Const.**
1	pen up on transition strokes not in base-shape	**1**	pen up on transition strokes not in base-shape
	closed curve start position determines motion direction		closed curve start position determines motion direction

	no R-L sequencing of base strokes		no transition strokes
2	no pen lifts	**2**	no D-U base strokes
3	no transition strokes		minor base strokes after major base strokes
4	no D-U base strokes	**3**	no pen lifts
	minor base strokes after major base strokes		no R-L base strokes
5	no R-L base strokes	**4**	no D-U sequencing of base strokes (v2)
	no D-U sequencing of base strokes (v2)	**5**	start at rightmost possible start point
	start with stroke containing point closest to upper left corner	**6**	start with stroke containing point closest to upper left corner
0	*start at rightmost possible start point* *	**0**	*no R-L sequencing of base strokes* *

We therefore ran FRed on the combined set of 11 constraints for the two participants, to attempt to reveal the necessary relationships that are similar and different between these two participants. We found that RYI and AAL only differed in two necessary domination relations: AAL needed ‘no pen lifts’ to dominate both ‘no transition strokes’ and ‘no down-to-up base strokes’, whereas RYI needed ‘no pen lifts’ to be dominated by the latter two constraints. Of course, AAL also needed ‘start at rightmost start point’ to dominate ‘start with stroke containing point closest to upper left corner’, whereas RYI did not need the first of those two constraints at all. For AAL, the constraint ‘start at rightmost start point’ was needed to explain the production of x and d starting at the top-right. However, AAL still produced most letters starting on the left (e.g., b, h, m, n, etc.), because of other higher-ranked constraints. For example, for lowercase m, starting on the right would also mean starting on the bottom, and moving rightward with the stroke, thus violating the higher-ranked constraints ‘no down-to-up strokes,’ and ‘no right-to-left strokes.’ In contrast to AAL, RYI needed the constraint ‘no right-to-left sequencing of strokes,’ which was not required for AAL at all, to dominate ‘no down-to-up sequencing of strokes.’ This domination relation was required to explain why RYI produced lowercase d starting with the curve, and producing the vertical line second.

Much like in the case of GEM and LLN, we needed only a small change to the necessary domination relations (two opposing pairs) to account for the differences in stroke patterns between RYI and AAL. But will there be a bigger number of opposing domination relations when we contrasted sets that differ by a lot more, for example the prescribed writing and actual participant writing? We compared the ranking of constraints for prescribed writing of Roman lowercase letters with the ranking for one of the participants who diverged the most from the prescribed stroke patterns. Participant ACG used the prescribed stroke pattern for only 14 of the 26 Roman lowercase letters (54%). And yet, despite a huge difference in stroke patterns, we saw only a minor difference in the constraints used.

Of the nine constraints needed to account for the prescribed writing of Roman lowercase, and the 10 needed to account for the writing of participant ACG, 8 constraints were shared (Table 9). Note that the two constraints not required for the modeling of the Roman prescribed, ‘no transition strokes’ and ‘pen up on transition strokes not in base-shape’, remain in the highest stratum and do not get demoted simply because they do not pick a non-target over a target, and not because they are required to dominate any other constraint. In fact, the ranking would work equally well if these two constraints were ranked in the lowest stratum.

Table 9. Ranking of constraints used to model the writing of participant ACG (left) and the prescribed way of writing Roman lowercase letters (right). Where a constraint has multiple versions, the version is indicated in parentheses. Where a constraint is not required to be in the minimal set, it is marked with italics and an asterisk. ACG shares 8 constraints with the Roman prescribed, and their constraints are ranked in a different order. Str. = stratum; Const. = constraint. See Table 20 in Appendix A for a complete description of each constraint and its variants.

	ACG		**ROMAN PRESCRIBED (LC)**
Str.	**Const.**	**Str.**	**Const.**
1	no transition strokes	**1**	*no transition strokes* *
	pen up on transition strokes not in base-shape		*pen up on transition strokes not in base-shape* *
	no R-L sequencing of base strokes (v1)		no R-L sequencing of base strokes (v1)
	minor base strokes after major base strokes (v3)		minor base strokes after major base strokes (v3)
	curves continue prior motion direction		curves continue prior motion direction
2	no pen lifts		pen up on all transition strokes
	no D-U base strokes	**2**	no pen lifts
3	no R-L base strokes		no D-U base strokes
	curves counter-clockwise	**3**	no R-L base strokes
	start with stroke containing point closest to upper left corner		curves counter-clockwise
	pen up on all transition strokes *		start with stroke containing point closest to upper left corner

Using FRed to compare ACG's ranking to the ranking for prescribed Roman lowercase, we found there was only one opposing domination relation between them: Roman prescribed needed 'pen up on all transition strokes' to dominate 'no pen lifts', and ACG needed the opposite domination relation. ACG's preference to avoid lifting the pen even at the cost of writing (pen-down) on transition strokes was mirrored with many other participants, in both the Hebrew and English groups. For example, in contrast to the prescribed stroke pattern, ACG preferred to repeat the straight vertical stroke in the letter 'p' rather than lift the pen to get to where the curved stroke began. However, ACG produced the letter 't' just like the prescribed stroke pattern – lifting the pen up to get to the horizontal crossing line.

The similarity in the stroke pattern for the letter ‘t’ highlights another aspect of the complex relationship between constraints, which reflects a widespread distinction between prescribed and actual writing: Whereas prescribed writing dictates a pen lift whenever a stroke begins at a different point from where the previous stroke ended (i.e., “pen up” on all transition strokes), actual writers tend to only lift the pen if the transition passes through points not in the character shape, and keep the pen down if they are merely repeating already written strokes. For example, whereas Hebrew-speakers always lifted the pen when transitioning to the left-most stroke in the Hebrew letter He (Figure 12a), most of them left the pen down when transitioning to the left-most stroke in the letter Het (Figure 12b), repeating the first stroke in the opposite direction. In fact, not a single participant, in either the Hebrew or English group, used the constraint ‘pen up on all transition strokes’; every single one of them opted instead for ‘pen up on transition strokes not in base-shape’. By only lifting the pen when a transition stroke would deviate from the base-shape, participants evidently prioritized speed (i.e., avoiding the time cost of lifting the pen) over legibility (by allowing a previously written stroke to be repeated).

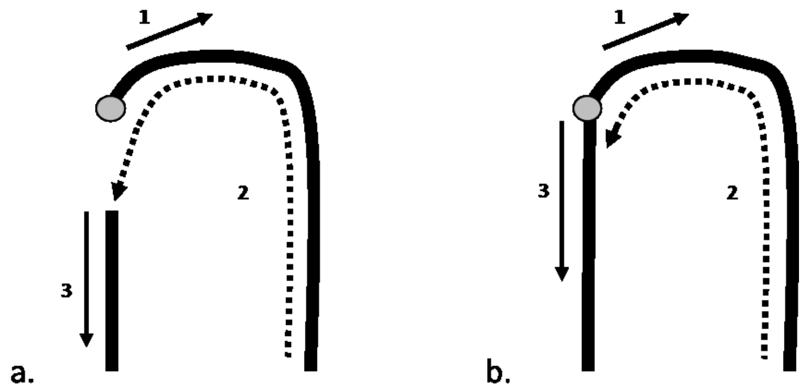


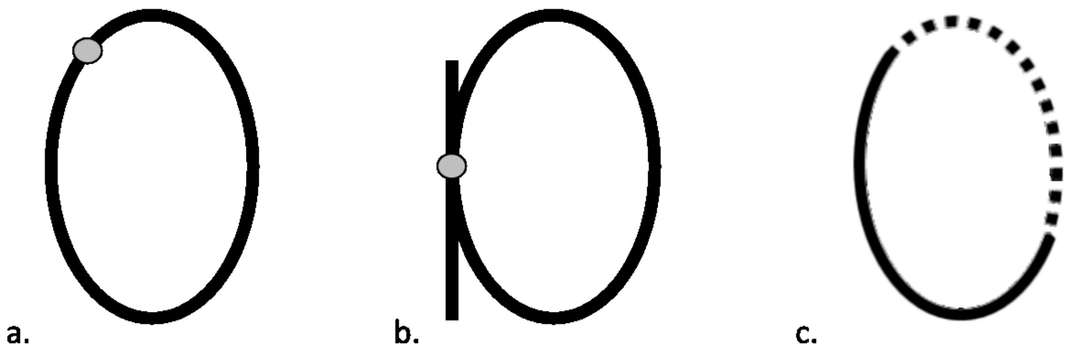
Figure 12. Hebrew script He (a), and Hebrew script Het (b). The gray dot marks the prescribed starting position, and the numbered arrows mark the prescribed stroke order and direction. The dotted line marks the transition stroke; whereas in He (a) this stroke is never produced pen-down, in Het (b) it often is.

Motion direction when producing circles

While analyzing the data, we also noticed that participants differed with respect to their preferred motion direction when producing circles. The vast majority of English speakers (22/24) produced the letter 'o' counter-clockwise (of the remaining two participants, one was left-handed). Hebrew speakers, on the other hand, were split on the production of closed curves, such as in the letter Samech (a circle very similar to the letter 'o'; Figure 13a). Nine of the Hebrew-speakers produced all closed curves clockwise; two produced all closed curves counter-clockwise; and eleven Hebrew speaking participants produced the closed curve in the letter Samech (Figure 13a) counter-clockwise and the closed curve in the final form of the letter Mem (Figure 13b) clockwise. The remaining two participants were inconsistent within letter, producing Samech both clockwise and counter-clockwise on different trials.

When we inspected all the produced occurrences of the Hebrew script Samech, we noticed that while almost all the participants produced it starting roughly in the top half of the circle (between 8:00 and 2:00, if imagining the letter as the face of a clock), there was a tendency towards one direction (clockwise vs. counter-clockwise) depending on whether the starting position was closer to the top-right part of the circle (within the dotted line in Figure 13c) or the left-bottom part of the circle (within the black line in Figure 13c).

Figure 13. Hebrew script Samech (a), Hebrew script Mem (word-final form, b). Starting position (c): Participants who produced Samech counter-clockwise tended to start in the top-right half (dotted line), and participants who progressed clockwise tended to start in the bottom-left half (full line). The gray dot in panels a. and b. represents the prescribed starting position.



Of the 14 participants who produced Hebrew script Samech counter-clockwise, 12 consistently began the circle from a point along the top-right half, whereas 10 of the 11 participants who produced it clockwise began the circle from a point along the left-bottom half. One participant produced Samech both clockwise and counterclockwise equally frequently, and this participant was also inconsistent with respect to her starting position and its relation to the direction of movement. In the case of final Mem, all but two of the participants started the circle near the leftmost point, in anticipation of the next stroke (a vertical line to the left of the circle), and indeed 21 out of 24 participants produced it clockwise. All three participants who produced it counter-clockwise also started from a point on the top-right arc of the circle.

These two findings regarding circle writing direction (in Hebrew script Samech and in Hebrew script final-Mem) are consistent with Van Sommers' findings (1984): He instructed participants to copy circles starting at given points along the circle's circumference, and noticed that participants preferred to progress clockwise when starting from the bottom-left half of the circle and counter-clockwise when starting from the top-right half of the circle (Van Sommers, 1984). Meulenbroek, Vinter, and Mounoud (1993) note that participants who start circle production at the top tend to go counter-clockwise, and those who start at the bottom tend to go clockwise. As Meulenbroek et al. note, several factors may be contributing to the direction of circle production. The position of the hand on the paper, and the angle of the pen-tip may facilitate counter-clockwise movement when starting near the top-right and clockwise movement when starting more to the left, allowing for a better-controlled movement as well as lower chance of ripping the paper. Meulenbroek et al. (1993) also suggest that these directions allow better visual monitoring of the movement.

The reasons behind people's preference for starting position may be slightly more obscure, although culture and instruction undoubtedly play a role in the decision. A recent article by Ha & Sonnad (2017) analyzed circle drawings from 40,000 participants worldwide, and notes that country of origin is highly correlated with the direction of production of the circle. Meulenbroek et al. (1993) also note that while the direction of circle production is somewhat dependent upon starting position in children, this tendency becomes much more pronounced in adults who have been taught how to write English. A similar finding for Japanese speakers' tendency to produce circles clockwise is reported by Taguchi and Noma (2005). Both Goodnow et al. (1973), and Amenomori, Kono, Fournier, and Winer (1997), compare circle direction across cultures. Goodnow et al. (1973) report that while American children and adults prefer to produce circles counter-clockwise, Israeli Hebrew-speaking children and adults prefer to produce them clockwise. Similarly, Amenomori et al. (1997) show that while a growing percentage of American children produce circles counter-clockwise as they grow up, Japanese children exhibit the opposite trend.

To account for the preference of stroke production in our data, we defined a new constraint. Rather than a constraint dictating the direction of stroke production independently from the starting position, we defined a constraint that dictates a clockwise production if the starting point is in the left-bottom half and counter-clockwise production if the starting point is in the top-right half. Using this constraint, we were able to explain not just why participants produce isolated closed curves in one direction or the other, but also why closed curves that follow or lead into another stroke are produced this way. For example, whereas the closed curve in lowercase d is almost always produced counter-clockwise (in accordance with its starting position on the right), the similar curve in lowercase b, which starts on the left, is usually produced clockwise, even by native English speakers.

Discussion

Since OT implements a strict relationship between constraints, one concern was that it might not be flexible enough to account for the variance observed in actual participants' handwriting (unlike prescribed writing). However, our results show a remarkable extent to which we can model participants' handwriting in terms of violable constraints ranked in strict domination. This was true for participants among whom we observed different writing styles and varying stroke patterns, and for two separate sets of characters, learned in two distinct parts of the world, that are written from opposite directions. However, we were not entirely successful in modeling the writing of all our participants. With 33% of the English-speakers, and 17% of the Hebrew-speakers, we achieved only partial success, meaning we were able to account for their writing only when omitting one target from the set. For another 33% of the English-speakers and 8% of Hebrew speakers we could not find a ranking that explained their writing at all. A failure to find a ranking that accounts for participants' handwriting within the framework of OT and our constraints could be the result of simple small idiosyncrasies in the way people write certain letters, or it could stem from some other limitation in the way we applied OT to handwriting. We discuss this point further in Chapter 4 and in the General Discussion.

Our considerable success in modeling most participants' handwriting suggests that the systematicity we observed in the produced stroke patterns can in fact be explained using violable constraints ranked in strict domination, thus giving a positive answer to one of the main questions of this work. Another goal was to understand the factors contributing to variability in stroke patterns among people. We asked whether the ranking of constraints can explain the differences among participants in terms of their produced stroke patterns. As we alluded to earlier, with a large-scale application of OT, comparisons between rankings of constraints are not trivial, and neither is drawing conclusions from those comparisons. For one thing, our ranking is not fully-determined, as the constraint demotion algorithm does not specify a ranking within the same stratum (resulting in multiple possible full rankings for almost every participant). Secondly, a few participants needed slightly different manifestations of the same constraint in their ranking, and so directly comparing them is not straightforward. The answers do not emerge on their own; part of our goal in this work was to define the methods for using the OT modeling framework, and additional analyses are required to interpret the results.

Although the rankings of constraints are not straightforward to interpret, we were able to compare the necessary domination relations between two sets of ranked constraints using FRed (Brasoveanu & Prince, 2005). We used this tool to detect where two constraint domination relations diverge, which informed us of the underlying reasons for a difference in stroke patterns. We used FRed to compare the necessary domination relations between participants who differ only in the stroke pattern used for one character (GEM and LLN), between participants who differ in their stroke patterns on 20% of characters (RYI and AAL), and between the prescribed way of writing Roman lowercase letters and one representative participant's actual production (ACG), which differed by nearly 50% of characters. In this chapter we made several illustrative analyses to prove the viability of the approach, although a systematic comparison between every pair of participants, or between each participant and the prescribed writing, is certainly possible.

We have found that differences in participants' stroke patterns can be traced back to differences in the underlying causal structures. Interestingly, we have found that both minor differences (diverging by one stroke pattern) and substantial differences (diverging by 20% or 50% of stroke patterns) can be caused by very subtle differences in the necessary domination relations. A major change in the overt manifestations can be explained with a seemingly minor tweak to the underlying causal structures. Interestingly, the opposing domination relations in the cases described above occurred in higher strata the more targets were different.

Other differences in handwriting style were evident in participants' minimal sets of constraints. For example, five Hebrew speaking participants used 'start at leftmost possible starting point', whereas eleven others used ‘start at rightmost possible starting point’. The difference between the two groups of participants is manifest in two letters: Ain (Figure 14a) and Aleph (Figure 14b). The participants who used ‘start at the left’ wrote Aleph starting with the vertical stroke on the left, and Ain starting at the top-left and going clockwise (e.g., Figure 14c). Participants who used ‘start at the right’ produced Aleph starting with the curve on the right, and Ain starting at the top-right and going counter-clockwise (Figure 14d). Six participants needed neither constraint to account for their handwriting, and all six produced Aleph starting at the top-right, and Ain starting at the top-left.

With the exception of one participant, who produced most of the letters from the bottom-up, and whose handwriting we could not model, all other participants in the Hebrew group, including those who used ‘start on the right,’ produced the vast majority of Hebrew script letters starting on the top-left, just like the Hebrew prescribed writing. This despite only five of the 22 participants whom we could model using the constraint ‘start on the left’. Just like for Hebrew prescribed, other constraints that dominated ‘start on the right’ dictated the production of letters starting on the left. For example, the constraint ‘no D-U strokes’ had to dominate ‘start on the right’ for all 11 Hebrew speakers who used the latter, thus preventing the production of letters such as Reish (ר, the second letter in Figure 14c and 14d) from starting at the bottom-right. Instead, the ranking of constraints (e.g., ‘no right-to-left strokes’ and ‘no pen lifts’) dictated that the preferred production start at the top-left.

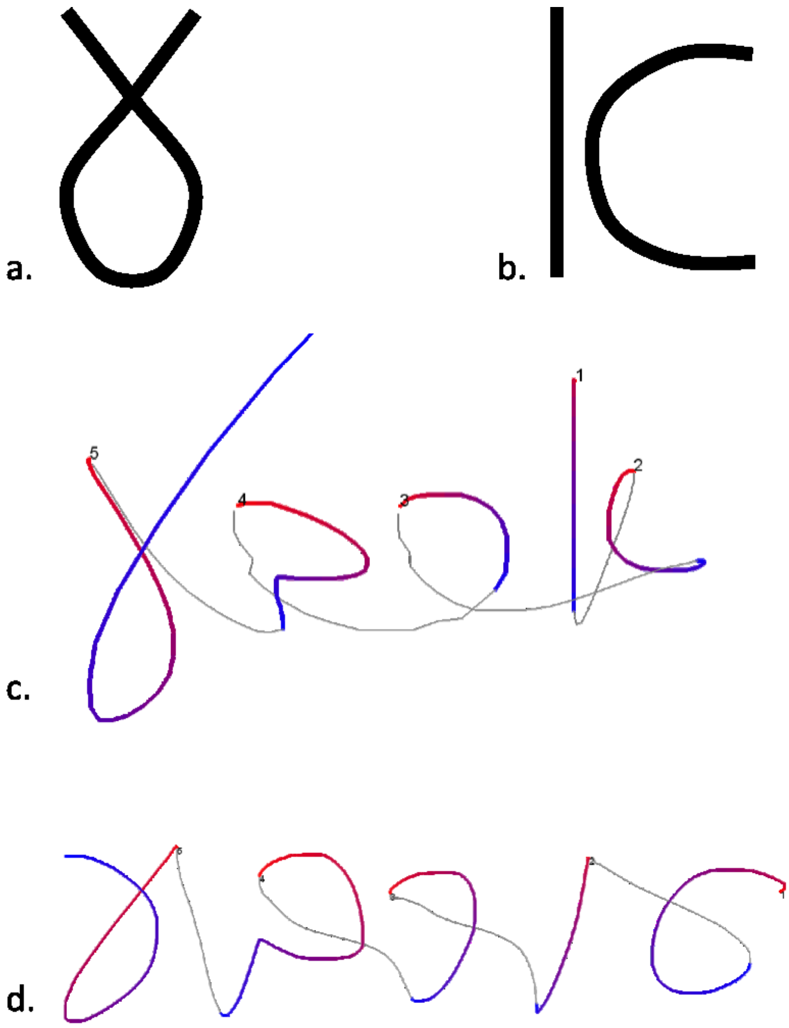


Figure 14. Top: The Hebrew script letters Ain (a) and Aleph (b). Bottom: Two participants' productions of the Hebrew word /arba/ (four), spelled (from right-to-left) Aleph, Reish, Beit, Ain. One participant produced both Aleph and Ain starting at the top-left (c), and another produced both Aleph and Ain starting at the top-right (d). The strokes' color represents direction of production, from red (beginning of the stroke) to blue (end of the stroke). Light gray lines represent pen-up movement.

Interestingly, the distinction of starting position also occurred in the prescribed data, with Hebrew print using 'start at leftmost possible starting point' and Hebrew script using 'start at rightmost possible starting point'. This seems to be related to the relative importance of legibility over speed. Starting at the leftmost possible point sacrifices speed, as it means starting farthest from where the previous letter ended. However, it positions the writer better to produce left-to-right strokes, consistent with better control of motion for right-handed individuals (Van Sommers, 1984), thus improving legibility. Starting at the rightmost possible starting point takes the opposite approach, favoring speed over legibility. In Hebrew print, while starting at the leftmost possible point was required for the model, it also had to be dominated by 'no left-to-right sequencing of strokes,' meaning that production only started on the left if it did not violate the right-to-left order of strokes (or if this direction complied with another higher-ranked constraint).

Actual writing of Hebrew print vs. Hebrew script

The production of Aleph and Ain above represents one interesting case in which some participants' writing deviates from the prescribed production. But Hebrew offers a unique opportunity to examine the difference not just between prescribed and actual writing, but between a character-set designated for reading and one designated for handwriting. One truly remarkable finding was that 12 out of the 24 Hebrew-speaking participants, all of whom live in Israel and are exposed to Hebrew print regularly and frequently, could not remember how to produce some Hebrew print letters in handwriting. One participant could not recall how to write 8 out of 27 letter-shapes. For comparison, not a single Hebrew-speaking participant had a similar problem with Roman letters, to which they are surely exposed less frequently on a daily basis.

Many of the Hebrew speakers produced Hebrew print letters with multiple different stroke patterns, and occasionally with variations on the basic shape (producing slightly different allographs of the same letter) throughout the testing session. Perhaps because of this variance within individuals' handwriting of Hebrew print letters, as well as the inconsistency stemming from participants' inability to recall the correct shapes, we were unable to model the handwriting of Hebrew print in any of our participants. This might suggest that handwriting of Hebrew print is not mastered as a rule-governed skill in the way that handwriting of Hebrew script is, or that the ranking of constraints has not been fully established for Hebrew print.

The effect of language context and instruction

Since we collected data from Hebrew speakers writing English, we were able to make a direct comparison of each participant's handwriting in different languages. Although a full analysis of these data is outside the scope of this dissertation, we do wish to point out some observations that may shed light on how constraints are applied. Some differences between Hebrew speakers' writing in English and in Hebrew were obviously due to the different shapes of letters in the two languages. However, at least seven Hebrew script letters share practically the same shape with Roman letters (see Figure 18 in Appendix B), but were often produced by the same participant using different stroke patterns, depending only on the language-context in which they were presented (Figure 15). This can be explained by considering that different rankings are operative for the two writing systems, and those give rise to different motor plans, even when considering the same shape and writing by the same individual.

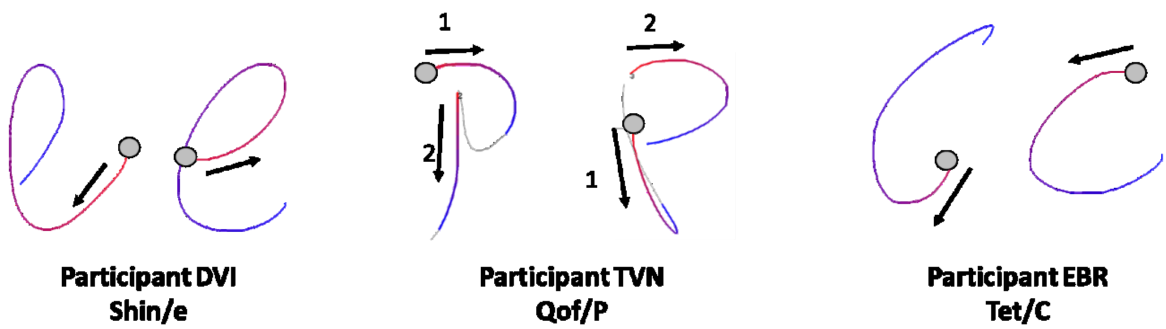


Figure 15. Left: Participant DVI producing Hebrew script Shin (left) from the right, and English lowercase e (right) from the center. Middle: Participant TVN producing Hebrew script Qof (left) starting from the top curved stroke, and English uppercase P (right) starting from the left straight stroke. Right: Participant EBR producing Hebrew script Tet (left) going clockwise, and English uppercase C (right) going counter-clockwise. Arrows represent stroke direction, the gray dot represents starting position, and the numbers represent stroke order.

The language context raises an interesting comparison to the case of Samech and final-Mem described earlier. In the case of the closed curves, one might speculate that the relation between directionality and starting point is tied, at least in part, to the angle of the pen on the paper, and the effort not to rip the paper. The differences in stroke patterns between Hebrew speakers and English speakers could therefore be related to their hand positions, and other individual differences. However, in the case of similar letter shapes in different languages, individual preferences and the presence of other strokes in the letter cannot be the primary motive driving the different stroke patterns. The reason must be tied to cultural and instructional biases that manifest only when producing some shape within the context in which it was learned.

Chapter 3: Participant writing – right and left-handed individuals

Like writing in different languages, the writing of people with different handedness preferences provides an opportunity to examine cases of conflicting principles. Around 10% of the population are thought to be left-handed (Hardyck & Petrinovich, 1977). Previous research on left-handed writing and drawing has examined hand posture and its effects on writing in right-handed (RH) and left-handed (LH) individuals (e.g., Meulenbroek & Van Galen, 1989; McKeever, 1979; Peters & McGrory, 1987; Wing, 1979). Other researchers report on the direction in which simple line drawings face (e.g., De Agostini & Chokron, 2002; Picard, 2011; Van Sommers, 1984), and the direction in which circles are produced (e.g., Van Sommers, 1984) when drawn by RH and LH individuals. Simner (1984), and others, have looked at the interaction between handedness and left-right reversal errors of letters by children who are just learning how to write. And yet, children are still taught to write letters using the stroke-patterns that are best suited for RH individuals. In this chapter, we examine the handwriting of 20 RH and 20 LH participants writing in English, and use it to illustrate how the conflict between good control of motion and overall writing direction is resolved.

Methods

Participants

Forty native English speakers (26 females, 14 males) participated in this experiment. None of the participants had any known language or neurological impairments, and none had participated in the previous experiments. Half of the participants were left-handed (LH), as determined by their self-reported writing-hand, and the other half were right-handed (RH). Participants gave written informed consent in accordance with the Johns Hopkins Institutional Review Board and were given course credit for participation. The RH participant group consisted of 14 females and 6 males with a mean age of 19.7 years, and the LH group consisted of 12 females and 8 males with a mean age of 26.1. This age difference was not significantly different between the groups, $t(38) = 1.958$, $p = 0.06$. Participants had a mean of 13 years of education for RHs and 14 for LHs, $t(38) = 1.731$, $p = 0.09$ (see summary in Table 10).

	RIGHT	**LEFT**	**TOTAL**
N	20	20	40
FEMALE/MALE	14/6	12/8	26/14
MEAN AGE (SD)	19;8 (1;0)	26;1 (13;9)	22;10 (7;4)
MEDIAN AGE	19;4	19;8	19;6
MEAN EDUCATION, YEARS (SD)	12.7 (1.0)	13.5 (1.8)	13.1 (1.4)
MEDIAN EDUCATION, YEARS	12	13	12.5

Table 10. Summary of demographic details for the right and left-handed participants in this analysis.

We administered the Edinburgh Handedness Inventory (EHI; Oldfield, 1971) to verify the handedness of each participant. Simple everyday tasks are scored according to the hand with which participants perform them (a positive score for the RH, a negative score for the LH). The scores range from 100 (completely RH) to -100 (completely LH). While none of the participants were classified as having the opposite laterality of their self-identified writing hand, the range of responses was fairly wide. Self-identified RH participants' scores on the EHI ranged from 30-100 (median 95), and LH's scores in our sample ranged from 0 to -90 (median -50). Two RH participants (with scores of 30 to 33) and six LH participants (with scores of 0 to -20) were classified in the "Middle" category, rather than one of the left or right-leaning deciles, a result which is in line with the distribution of the general population (Oldfield, 1971). We have opted to still treat those participants as belonging to their respective groups, due to their preference for handwriting effector, in accordance with Corey, Hurley, and Foundas (2001).

Stimuli, procedure, and data analysis

Stimuli consisted of 53 mono-morphemic English words (average length 6.2 letters, SD 1.4, median 6). The list was a version of the list of words described in Chapter 2, but it was shortened to allow two administrations of the list within a one-hour time frame. Each letter appeared on the list a minimum of 4 times (j, q, z) and a maximum of 35 times (e), with at least two appearances in initial position and at least two in non-initial position. The list of words was dictated to each participant twice (in different word-orders, counter-balanced across participants), and they were instructed to repeat each word, and then write it in all lowercase letters. Each participant wrote one list entirely with their dominant hand, and one entirely with their non-dominant hand (the order of dominant and non-dominant hand was counterbalanced across participants). Data collection and experimental procedure were the same as described in Chapter 2.

Candidate stroke patterns were the same as those used in Chapter 1. The set of constraints used in this analysis was the same as in Chapters 1 and 2. Targets were identified for each participant as described in Chapter 2, and were defined separately for each participant for the dominant and non-dominant hand. The current analysis focuses on the dominant-hand data only, although we discuss possible directions for further investigation using the non-dominant hand data in the General Discussion. Data analysis was carried out as in Chapter 2.

Results

Different strokes for different folks

To better understand the differences between RH and LH participants, we first examined every participant's stroke patterns. On most letters, RHs and LHs produced much the same stroke patterns. More than 90% of each group (right and left-handers) produced the same stroke pattern on 12 character-shapes: a, c, e, g, h, l, q, s, u, v, w, and on the z allograph without the horizontal crossing line. On six other letters (r, d, i, j, k, and n) more than one stroke pattern was prevalent, but the stroke patterns were produced at similar rates among RH and LH participants. For example, 75% of RHs and 80% of LHs produced the letter r starting at the top left, and an additional 20% of each group produced r starting from the bottom left. The remaining RH participant produced r from the top-left, but lifted the pen between the two strokes, rather than repeating the vertical line (see Table 11 for summary).

Table 11. Agreement between right-handed and left-handed participants in stroke patterns produced for each of the Roman lowercase letters. Same indicates the participants produced each stroke pattern with similar rates. Slight difference indicates RH and LH participants differed in the stroke patterns they used, but that difference did not reach significance. Different indicates the two groups differed significantly in their rate of production of each stroke pattern.

LETTER	RH-LH AGREEMENT	DETAILS
a	same	>90% of each group produced same stroke pattern
b	slight difference	15% of RH produced the curve counter-clockwise, compared with 0% of LH
c	same	>90% of each group produced same stroke pattern
d	same	same rate of production for different stroke patterns
e	same	>90% of each group produced same stroke pattern
f	different	100% of RH produced the horizontal stroke from L-R compared with 15% of LH
g	same	>90% of each group produced same stroke pattern
h	same	>90% of each group produced same stroke pattern
i	same	same rate of production for different stroke patterns
j	same	same rate of production for different stroke patterns
k	same	same rate of production for different stroke patterns
l	same	>90% of each group produced same stroke pattern
m	slight difference	5% of RH started at the bottom left, compared with 20% of LH
n	same	same rate of production for different stroke patterns
o	different	100% of RH produced the circle counter-clockwise, compared with 75% of LH
p	slight difference	35% of RH lifted the pen when transitioning to the curve, compared with 15% of LH
q	same	>90% of each group produced same stroke pattern
r	same	same rate of production for different stroke patterns
s	same	>90% of each group produced same stroke pattern

t	different	100% of RH produced the horizontal stroke from L-R compared with 15% of LH
u	same	>90% of each group produced same stroke pattern
v	same	>90% of each group produced same stroke pattern
w	same	>90% of each group produced same stroke pattern
x	slight difference	80% of RH started at the top left, compared with 50% of LH
y	different	100% of RH started at the top left compared with 75% of LH
z	same	>90% of each group produced same stroke pattern
z (with crossing line)	different	100% of RH produced the horizontal stroke from L-R compared with 25% of LH

Other letters showed differences between the two handedness groups. For example, the two groups differed on their preferred direction for closed curves. Whereas 100% of RHs produced 'o' counter-clockwise, only 75% of LHs produced it in this direction, and the others produced it clockwise (χ^2 = 5.0, p = 0.026). Interestingly, while all of the RHs began writing the 'o' at or near the top of the circle, the LHs showed greater variability in their starting position, eight of them starting the circle at least on one occasion from a point along the bottom-left arc of the circle (approximately between 6 and 10 on the face of a clock). Of those participants, the three who started only to the left of the top, and never at the top, also consistently produced the circle in 'o' clockwise. The other five participants started their production of 'o' either at the top or to the left, and always progressed counter-clockwise, regardless of starting position.

The difference in preferred circle direction was also reflected in the stroke patterns used to produce the letter 'b', for which 15% of RHs produced the curve counter-clockwise (i.e., continuing the downward vertical stroke's direction) compared with 0% of the LHs, and to a lesser degree in the letter q, where 100% of RHs but only 90% of LHs produced the curve counter-clockwise. Notably, LH's tendency to produce circles clockwise was less pronounced than that tendency in RH Hebrew-speakers. This could possibly mean that the direction of circle production stems from preferences regarding stroke order (as opposed to stroke direction), but we think it is more likely to simply reflect biases due to instruction (Hebrew-speakers are taught to produce circles clockwise, whereas English-speakers, regardless of handedness, are most likely taught to produce them counter-clockwise).

RHs and LHs also differed in their starting position on the letters x and y. Whereas RH participants were overwhelmingly likely to start both letters at the top-left (80% for x and 100% for y), LH participants were much more evenly divided, only 50% of them preferring to start x at the top-left, and 75% starting y at the top-left. Both of these differences were significant between the groups ($\chi^2 = 3.9$, $p = 0.05$ for x and $\chi^2 = 5.0$, $p = 0.026$ for y). Other minor differences between the two groups included LH participants being slightly more likely to start the letter m at the bottom than RH participants (20% for LHs vs. 5% for RHs), and LH participants being slightly less likely to lift up the pen when transitioning from the vertical line in the letter p to the curve (15% for LH vs. 35% for RH). Neither of the latter two comparisons reached significance ($\chi^2 = 2.0$, $p = 0.16$, and $\chi^2 = 2.1$, $p = 0.15$, respectively).

The starkest difference between RH and LH participants was in the direction of production of horizontal strokes. The lowercase Roman letters that have straight horizontal strokes are f, t, and the allograph of z with a crossing line in the middle. All of the RH participants produced the horizontal crossing lines in f, t, and z from left-to-right (that includes each of the 20 participants for f and t, and each of the 7 participants who produced this particular allograph of z). In contrast, 15 of 20 LH participants produced the horizontal line in both f, t, and z from right-to-left (including 3 of the 4 participants who produced the z allograph with the crossing line), another one produced the horizontal line in f from right-to-left but in t from left-to-right, and two others produced the horizontal crossing lines from different directions on different trials. Only two LHs produced horizontal crossing lines consistently from left-to-right (see Table 12 for a summary). The different rate of production of the stroke from left-to-right in RH and LH individuals was statistically significant for both f ($\chi 2$ = 28.7, $p < 0.001$), t ($\chi 2$ = 27.6, $p < 0.001$), and z ($\chi 2$ = 6.5, p = 0.01).

HANDEDNESS	LETTER	LEFT-TO-RIGHT	RIGHT-TO-LEFT	MIX	N
RH	f	100%	0%	0%	20
LH		15%	85%	0%	20
RH	t	100%	0%	0%	20
LH		15%	75%	10%	20
RH	z (with crossing line)	100%	0%	0%	7
LH		25%	75%	0%	4

Table 12. The direction of production for the horizontal stroke in lowercase f, t, and the allograph of z in which there is a middle crossing line, for right-handed and left-handed participants. RH = right-handed participants; LH = left-handed participants, N = the total number of participants represented in each row.

OT modeling of RH and LH writing

One of the main questions of this dissertation was whether the framework of ranked violable constraints is useful in shedding light on the mechanisms underlying writing. In this chapter we ask specifically whether this framework can illuminate the selection of stroke patterns for production when different (and often competing) principles are involved. To find out how conflicts between principles are resolved, we attempted to model the handwriting of our RH and LH participants using OT. We found a consistent ranking of constraints that fully accounted for all and only target stroke patterns for 10 of the 20 RHs and 10 of the 20 LHs. We had partial success (i.e., a ranking that accounts for all but one target) with an additional 6 RHs and 3 LHs (we discuss further the cases of failure to model in Chapter 4 and in the General Discussion). While the success of our modeling efforts was not complete, the similar level of success with RH and LH individuals points to our ability to tap the underlying representations governing handwritten production to a similar degree in the two groups, providing evidence supporting our interpretation of a joint causal structure underlying stroke production.

The main difference between RHs and LHs is the direction of movement that is best controlled. Whereas RHs have better control of motion (and thus increased accuracy and speed) when producing horizontal strokes from left-to-right, LHs have better control of motion when producing horizontal strokes in the opposite direction – from right-to-left. The direction of best-controlled movement interacts with the principle of writing in the direction that conforms to the overall direction of reading and writing (i.e., the direction of transition between letters). In the case of English, the overall direction of writing is from left-to-right. While in RH individuals the two principles result in the same outcome – a movement from left-to-right – for LH individuals those principles clash. For them, writing horizontal strokes from right-to-left follows good control of motion whereas writing from left-to-right follows the overall writing direction.

As in the case of Hebrew, described in Chapter 1, in which there was a conflict between these two principles (writing in the direction of best-controlled movement and writing in the direction of overall progression between letters), here too, as revealed by our modeling, the conflict was not resolved by simply ignoring one of the principles. The constraints needed to model prescribed (i.e., right-handed) Hebrew writing included both 'no right-to-left direction of strokes' and 'no left-to-right sequencing (or order) of strokes'. For LH English speakers, the constraints included both 'no left-to-right strokes' and 'no right-to-left sequencing of strokes'. Thus, each ranking adheres to both the overall writing direction of the language and the direction of stroke production in which movement is best controlled (see Table 13 for a summary of all the constraints used by RH and LH participants).

No other difference in constraint usage in the minimal set was as evident as the difference between using ‘no right-to-left base strokes’ and using ‘no left-to-right base strokes’. In fact, almost all other constraints were used at about the same rate among participants from either group. Importantly, constraints grounded in good control of motion and in writing in a direction that conforms with the overall direction of reading and writing applied similarly to RH and LH participants, as long as the direction of best-controlled movement was not dependent on the effector (Table 13). For example, both the RH and LH participants all needed the constraint ‘no down-to-up strokes’ to be ranked within their minimal set (that is, a successful modeling could not be achieved without it), because writing from top to bottom is a better-controlled motion regardless of the hand used to produce the stroke (except in the case of inverted hand posture, which we discuss later). Similarly, since participants in this analysis were writing English words only, the overall direction of reading and writing was left-to-right for all of them, and so again, every single one of the participants we could model needed the constraint ‘no right-to-left sequencing of base strokes.’

Table 13. The number of participants who used each constraint in their minimal set (out of 10 RH and 10 LH participants). RH = right-handed participants; LH = left-handed participants.

CONSTRAINT	**RH**	**LH**
no pen lifts	10	10
no transition strokes	10	10
pen up on transition strokes not in base-shape	10	10
no right-to-left sequencing of base strokes	10	10
minor base strokes after major base strokes	10	10
no down-to-up base strokes	10	9
closed curve start position determines motion direction	10	8
start at leftmost possible start point	8	7

no right-to-left base strokes	10	2
no left-to-right base strokes	0	8
first base stroke starts between baseline & x line	4	3
no down-to-up vertical base strokes	2	2
transition stroke goes to closer end of next stroke	2	1
start with stroke containing point closest to upper right corner	2	1
no initial down-to-up base strokes	1	3
no down-to-up sequencing of base stroke start position	1	1
curves counter-clockwise	0	2
curves continue prior motion direction	0	2

In contrast, 10/10 RH participants needed the constraint 'no right-to-left base strokes' compared with only 2/10 LHs. The other eight LH participants used 'no left-to-right base strokes', which was obviously not needed for any of the RH participants. The two LH participants who used 'no right-to-left base strokes' instead of 'no left-to-right base strokes' were the same two who produced the horizontal crossing lines in f, t, and z from left-to-right, like the RH participants. One of the two LHs who produced horizontal strokes left-to-right, CKH, had a handedness score of 0 (on the Edinburgh Handedness Inventory), indicating she uses her right and left hands equally often in everyday tasks. The other participant, HLS, had a handedness score of -90, placing him in the top decile of left-handers in terms of laterality.

As we mentioned earlier in this chapter, LHs' slightly greater tendency to produce curves clockwise was almost always accompanied by a difference in starting position compared to RHs: LHs who produced the letter 'o' clockwise started the production from a point to the left of the arc rather than from the top. Thus, our constraint 'closed curve start position determines motion direction' (described in Chapter 2), applied to their writing. Some cases that deviated from this pattern could be explained using other constraints (e.g., some constraints needed the constraint ‘curves counter-clockwise’ instead of ‘circle direction depends on starting position’), but others could not be modeled at all. Other differences in stroke patterns, such as RHs’ greater tendency to lift the pen when writing p, or LHs’ greater tendency to begin x and y at the top right as opposed to the top left, were reflected in the ranking of constraints for individual participants. For example, participant CSG, a LH who produced the letter x starting at the top-right, needed the constraint ‘start with stroke containing point closest to upper-right corner’ to be ranked, although he still needed ‘no right-to-left sequencing of strokes’ to dominate it to account for other letters being produced from the left.

Discussion

We found systematicity in the handwriting of RH and LH participants, just as we did in the handwriting of Hebrew and English speakers described in Chapter 2, and were able to model this systematicity within the framework of ranked violable constraints to some extent. Our methods were useful to illuminate some of the key differences between the stroke patterns most commonly used by RH and LH participants. Some of the differences between right and left-handers, such as the direction of production of horizontal strokes, were easily traced back to the direction that produces an abduction movement, which could be considered less costly and more accurate (e.g., Bradshaw, Bradshaw, & Nettleton, 1990). The reasons behind other differences in stroke patterns, such as LHs’ greater tendency than RHs to produce circles clockwise are slightly less clear, although they could still be tied to the direction of best-controlled movement.

The direction of circle production could also be tied to another set of differences between RHs and LHs, concerning LHs’ greater tendency to produce strokes from the bottom up. This difference between RHs and LHs was particularly evident for the letter m, although we observed this pattern with other letters to a lesser degree (e.g., i, x). Perhaps relatedly, RHs were more likely to lift the pen up on the letter p, thus avoiding a repeat of the vertical stroke, but also avoiding producing a bottom-up stroke. The reason behind LHs’ greater tolerance to bottom-up strokes could be related to the angle in which the pen touches the paper. If the tip of the pen points downwards rather than upwards, a vertical stroke could be better controlled going up instead of down.

The phenomenon in which the individual holds the hand above the x-line (the line of writing), with the tip of the pen slanted toward the bottom of the page is called Inverted Hand Posture (IHP; e.g., McKeever, 1979; Meulenbroek & Van Galen, 1989; Peters & McGrory, 1987). IHP is much more common in LH than in RH individuals, occurring in about 30-70% of LHs compared with 1-10% of RHs. Wing (1979) notes that flexion of fingers is usually used to control motion in the vertical plane (i.e., letter height), but radial and ulnar abduction of the wrist is used to control motion in the horizontal plane (i.e., letter width). However, in individuals with IHP, the two systems swap functions, with the wrist used to control movement in the up-down plane and the fingers used to control movement in the left-right plane. Therefore, in LH individuals with IHP, movement from the bottom up may be better controlled than movement from the top down, as it would constitute abduction movement for the wrist. We only asked the experimenters collecting data for our experiment to note hand posture if they saw “anything unusual”, and none of them reported seeing that. However, if IHP indeed occurs in 30-70% of the LH population, undoubtedly some of our participants have exhibited it, and our experimenters either did not find it “unusual”, or the differences in posture were subtle enough not to be noticed. In future experiments we plan to record hand posture as well as the angle of the pen on the paper, to better understand the effect of those factors on stroke patterns.

Our modeling with OT proved successful to the same degree for both RH and LH participants. We found that the ranking of constraints was useful as a tool to highlight the differences between the groups, as well as the similarities between them. Most importantly, we have found that whereas both RH and LH participants sequenced letter-strokes from left-to-right, following the overall direction of reading and writing in English, only the LH group produced individual strokes from right-to-left, as evidenced by their exclusive use of the constraint ‘no left-to-right base strokes.’ Our level of success in finding a ranking of constraints that would account for all and only target stroke patterns was similar to (and even slightly greater than) the one we had with the English-speaking participants reported in Chapter 2, although lower than the level of success we had with the Hebrew speakers. As in Chapter 2, in this chapter too there were quite a few cases in which we only achieved partial success (i.e., a ranking that accounts for all but one letter's stroke pattern) or failed in finding a ranking.

Many factors could be contributing to our failure to model some participants' handwriting. We might not have identified all of the needed constraints, or possibly the way we implemented the framework does not capture the variability in stroke patterns. It could also be that participants' (or at least some participants') handwriting does not follow rules at all, in which case our efforts would never yield a successful model. But the systematicity we observed in handwriting across the board, and the fact that the vast majority of stroke patterns are never observed, lead us to believe that this is not the case. A final alternative, is that participants do follow violable constraints, but the way the underlying structure is organized is not with strict domination, but with another kind of relationship between the rules. We explore this possibility further in Chapter 4.

Chapter 4: Harmonic Grammar analysis of handwriting

In this chapter, we used Harmonic Grammar (HG; Legendre et al., 1990a), a sister-framework to OT, to model stroke patterns in writing. HG, like OT, deals with complex relationships between violable constraints. But unlike OT, HG does not require that the constraints be ranked in strict domination. Instead, the Harmony of each candidate is computed using a weighted sum of constraint violations. HG allows lower-ranked constraints to "gang-up" on higher-ranked ones, and thus a candidate which is less Harmonic according to OT, may nevertheless be chosen for production with HG. We applied the HG framework to the prescribed writing of Roman lower and uppercase letters, Hebrew print, and Hebrew script, as well as to the writing of the right-handed (RH) and left-handed (LH) participants described in Chapter 3.

We asked whether using the same constraints and only relaxing the requirement of strict domination would allow us to: 1) Explain the prescribed writing using fewer constraints (i.e., a smaller minimal set) than with OT; and 2) Model more participants' handwriting (i.e., explain the writing of some participants for whom we could not find a ranking using OT). We further used the framework of HG to more easily investigate the amount of work each constraint does in accounting for different letters' stroke patterns.

Methods

Participants and targets

In this analysis, we modeled the prescribed writing of Roman lower and uppercase, Hebrew print, and Hebrew script, as well as the handwriting of the 40 native English speakers described in Chapter 3. Target stroke patterns for the prescribed writing were identified as described in Chapter 1. Target stroke patterns for the participants were those identified for writing with the dominant-hand for each of the 20 RH and 20 LH individuals, as described in Chapter 3. Candidate stroke patterns, the set of constraints, and the computation of constraint violations were the same as those used in each of the OT analyses in Chapters 1, 2, and 3.

Harmonic Grammar data analysis

Unlike the analyses in the previous chapters, in this chapter we used Harmonic Grammar (HG) to model stroke patterns in writing. We computed each candidate's Harmony, equal to the negative sum of weighted constraint violations, and the candidate with the greatest Harmony of all the candidates for a specific character was declared the most Harmonic, or the optimal candidate, and the winner. In our implementation, if two candidates tied for the greatest Harmony, they were both declared winners. As with OT, when two candidates were designated as targets, both had to be more Harmonic than all other candidates, but we did not require that the two targets tie for the greatest Harmony, or specify that one should be more Harmonic than the other.

The Harmony for each candidate corresponds to a weighted sum of constraint violations. In the simplified example below (Table 14), candidate 1's vector of constraint violations is [1, 1], and candidate 2's vector is [0, 3]. The weight for constraint 1 is 1.0 and for constraint 2 is 0.4. To get each candidate's Harmony we multiply the number of violations of a given constraint by that constraint's weight, and sum over all constraints. We then subtract this weighted sum from 0, as each violation incurs a penalty, rather than a bonus. Thus, for candidate 1 the Harmony is:

$$0 - (1 * 1.0 + 1 * 0.4) = -1.4$$

and for candidate 2 it is:

$$0 - (0 * 1.0 + 3 * 0.4) = -1.2$$

You will notice that in the example below (Table 14) the target candidate (candidate 1) has lower Harmony than a non-target candidate, and thus it would not be chosen as the winner. In order for all and only target stroke patterns to be chosen as winners we needed to implement a gradual learning algorithm (GLA) to search for an optimal set of weights.

	CONSTRAINT 1	**CONSTRAINT 2**	**HARMONY**
CANDIDATE 1 (TARGET)	1	1	-1.4
CANDIDATE 2	0	3	-1.2
WEIGHT	1.0	0.4	

Table 14. Constraint violations and weights for two hypothetical candidates, and the resulting Harmony for each of them.

We implemented the GLA using Python, and ran it separately on each set of targets (e.g., one participant's handwriting). The GLA was initialized with a vector of random weights corresponding to the number of constraints. Each initial weight was a pseudo-random number between 0 and 1 drawn from a uniform distribution using Python's "random" package. The random number generator was initialized using a seed integer between 1-1000, chosen at random using the same package, and the seed was recorded for each participant for reproducibility. We used the initial weights to compute the Harmony of each candidate, and count how many of the targets were chosen correctly as the winners. If all and only targets were chosen as winners, the set of weights was deemed optimal, and we stopped.

However, if the set of winners was different from the set of targets (that is, at least one target was not chosen as the winner, or at least one non-target was chosen as the winner), we searched for a set of weights that better fit the data, using the following procedure. We sampled a letter at random (from the list of 26 Roman lowercase characters), using the same random number generator described earlier. We computed the Harmony for each of the letter's candidates by summing over the product of each constraint's weight and its violations. We then checked which candidate was chosen as the optimal for that character. If the target (or targets, in case some participant produced more than one stroke pattern for a given letter) was chosen as the most Harmonic, no further action was taken on that iteration, and we repeated the process again, sampling a new letter at random, until all and only targets were chosen. However, if some non-target candidate was more Harmonic than any of the targets for that character, we updated the weights.

To update the weights, we subtracted the number of violations of each constraint in the target candidate from the number of violations of that constraint in the "error" winner (the non-target that was more Harmonic than the real target). We multiplied each difference by a learning rate (in our model, set at 0.1), and added that to the vector of weights. In the example in Table 14 (above), since the target (candidate 1) is less Harmonic than candidate 2 (it has a Harmony of -1.4 compared with a Harmony of -1.2 for the non-target candidate 2), the weights will be updated. The new weights for each constraint will be

$$W_{new} = W_{old} - (V_t - V_e) * lr$$

where W_{old} is the old (or initial) weight given to the constraints, V_t/V_e is the number of violations of each constraint for the target or error (respectively), and lr is the learning rate. In the case describe above (Table 14, and see Table 15 for the updated weights), assuming a learning rate of 0.1, the new weight for constraint 1 will be:

$$1.0 - (1 - 0) * 0.1 = 0.9$$

and the new weight for constraint 2 will be:

$$0.4 - (1 - 3) * 0.1 = 0.6$$

The new Harmony for candidate 1 will then be:

$$H_1 = 0 - (1 * 0.9 + 1 * 0.6) = -1.5$$

and the new Harmony of candidate 2 will be:

$$H_2 = 0 - (0 * 0.9 + 3 * 0.6) = -1.8$$

making candidate 1 the more Harmonic candidate after the update (Table 15).

	CONSTRAINT 1	**CONSTRAINT 2**	**INITIAL HARMONY**	**HARMONY AFTER UPDATE**
CANDIDATE 1 (TARGET)	1	1	-1.4	-1.5
CANDIDATE 2	0	3	-1.2	-1.8
INITIAL WEIGHT	1	0.4		
CHANGE	(1 - 0) * 0.1 = 0.1	(1 - 3) * 0.1 = -0.2		
NEW WEIGHT	1 – 0.1 = 0.9	0.4 – (-0.2) = 0.6		

Table 15. Constraint violations for two hypothetical candidates, the initial constraint weights and Harmony for the two candidates, and the weights and Harmony following an update of the weights.

We continued sampling and updating the weights for a maximum of 10000 iterations or until all targets were chosen over all non-targets for each letter. Occasionally, our algorithm got “stuck” in a local minimum and could not find a set of weights that accounts for the entire data set even though such a set existed. To combat this problem, when the algorithm failed to find weights for the entire set of characters, we ran it again using new randomly initialized weights, up to three times. We report the best level of success we were able to reach as a percentage of letters for which the target was chosen out of the total number of letters modeled for each set of targets.

While many researchers limit the weights to only non-negative numbers, arguing that negative weights translate to an advantage rather than a penalty for violating a constraint (e.g., Potts et al., 2010; Prince, 2002), we have opted not to impose such a limitation. Our reasoning was twofold: First, each of the constraints we defined can fairly easily be translated to the opposite without rendering it meaningless. Second, by allowing negative weights we were able to pinpoint whether any participant is using a constraint in a way we did not anticipate. For example, as mentioned earlier, some participants (in particular LHs, although some RHs show this pattern too) hold their hands in an inverted handwriting posture, which makes strokes written from the bottom-up better controlled than strokes written from the top-down. Since we do not have the constraint ‘no up-to-down strokes’ in our set of constraints, allowing ‘no down-to-up strokes’ to receive a negative weight (and thus carrying an advantage when violated rather than a disadvantage) could illuminate the usefulness of the constraint we are missing. However, to adhere to standards set by previous researchers, we have also run our GLA with the limitation of no negative weights, and we report those results as well.

Results

Prescribed writing

We searched for a set of weights that would yield the greatest Harmony within the framework of HG for all and only target stroke patterns compared to the non-target alternatives for each letter. As expected, we found three sets of weights corresponding to the minimal constraint sets for prescribed writing of English, Hebrew print, and Hebrew script. One of the questions we wanted answered was whether the weights found in HG behave in a way that is similar to the strict domination found in OT. But comparing the ranking of constraints in OT and their weights in HG, and even interpreting the weights in HG on their own, proved less straightforward than we assumed. When we initialized the weights in HG to random numbers between 0 and 1, we saw that the final weights changed considerably from one run to the next, often causing a significant change in the relative "ranking" of the constraints as well (i.e., when ranking the constraints by their weight). For example, the constraint 'closed curve start position determines motion direction' was ranked second, with the second-highest weight, on one run of the GLA, but 11^{th} and second-to-last on another run, using the same data and only different initialization of the weights.

The differences in weights between different runs of the model serve to illustrate one of the challenges we faced with HG: That the weights, while seemingly intuitive to understand, are in fact tricky to interpret. A high weight for a given constraint could indicate that this constraint was very important, and needed to be weighted heavily, or it could indicate that this constraint was never updated. A constraint that is not causing any harm (i.e., not choosing any non-targets over a target) will never get updated, and remain with its initial weight – as low or as high as it was set. If the weights are initialized randomly (i.e., with different weights for different constraints), some of the constraints will show higher weights simply because they started high and did not change much. To investigate further the variability in the ranking of the weights, we ran the GLA on the Roman prescribed writing four times with randomly-initialized weights, and 4 times with all weights initialized to 1 (Table 16), and compared the ranking of constraints in each run.

CONSTRAINT	INITIALIZED TO RANDOM				INITIALIZED TO 1			
	R1	R2	R3	R4	U1	U2	U3	U4
pen up on all transition strokes	2	3	3	2	3	3	3	3
no pen lifts	6	8	6	5	6	6	6	6
no right-to-left base strokes	11	11	11	10	10	10	10	10
no down-to-up base strokes	9	10	9	7	7	8	8	7
no initial down-to-up base strokes	5	4	2	3	2	2	2	2
start at leftmost possible start point	7	12	8	9	11	11	11	11
no right-to-left sequencing of base strokes	10	6	7	8	7	8	8	7
no down-to-up sequencing of base strokes	8	5	4	5	7	6	6	7
minor base strokes after major base strokes	1	1	1	1	1	1	1	1
no down-to-up curved base strokes	3	7	5	4	4	4	4	4
start with stroke containing point closest to upper left corner	12	9	12	11	12	12	12	12
closed curve start position determines motion direction	4	2	9	11	4	4	4	4

Table 16. The minimal set of constraints needed for OT modeling of Roman lower and uppercase writing, and the relative ranking of the weights on each of 4 HG runs, when initialized to a random number between 0-1 (R1-R4), and when initialized to 1 (U1-U4).

As can be seen from Table 16 above, when weights were initialized randomly (R1-R4), the ranking occasionally varied widely (e.g., from 2nd to 11th for 'circle direction' constraint, or from 7th to 12th for 'start on the left'). However, it remained almost completely constant (for the same constraint on different runs) when all weights were initialized to 1 (U1-U4). When weights were initialized to 1, the only remaining random element in our algorithm was the choice of the next letter to sample. The differences in the ranking when initialized to random are therefore more likely to reflect the relatively small changes to some weights from their initialized value, rather than some required relations between constraints.

The weights that would change the least are those which, in our set, have the fewest conflicts with other constraints. For example, the constraint regarding ‘circle direction,’ which changes the most in the random initialization, does not come into conflict with any other constraint (and accordingly, is not required to stand in any domination relation with the other constraints in OT). Since it can only be violated once by any Roman character (no Roman character has more than one closed curve), its weight only needs to be greater than zero to distinguish production of the circle in the appropriate direction from the opposite direction. Indeed, the weight for this constraint never changed from its initial (non-negative) weight in any of our GLA runs, whether the weights were initialized randomly or uniformly to 1.

The inconsistent rankings (as well as raw weights) when initializing the weights randomly could mean two things. The first possibility is that constraints can have any weight within a certain range, and those weights can vary independently of one another within that range while still successfully choosing all and only target stroke patterns. However, this seems unlikely. For one thing, weights that were randomly initialized were still highly correlated with weights initialized to 1, even as the relative rankings of some constraints occasionally changed considerably ($r = 0.94$, $p < 0.001$). In addition, the average change in weights was also highly correlated in the random and the uniform initialization ($r = 0.86$, $p < 0.001$). The second and more likely option, which is also more in line with OT, is that some contingencies among the constraints exist, and dictate the relations between those constraints’ weights. A full analysis of the relation between different constraints’ weights in HG is certainly possible, but outside the scope of this dissertation.

Table 17. The minimal set of constraints needed for OT modeling of Roman lower and uppercase (top), Hebrew print (middle) and Hebrew script (bottom). For each constraint we list the raw weights associated with it in four runs of the HG GLA (W1-W4) when initialized to 1, its average HG rank, and its OT rank. Note that the constraint 'no down-to-up vertical base strokes,' which was required for a successful OT model, was not required to achieve full success with HG, and so was excluded.

CONSTRAINT	W1	W2	W3	W4	HG RANK	OT RANK
ROMAN UPPER AND LOWERCASE						
pen up on all transition strokes	1.1	1.2	1.2	1.1	3	1
no initial down-to-up base strokes	1.4	1.4	1.3	1.4	2	1
no down-to-up curved base strokes	1.0	1.0	1.0	1.0	4	1
minor base strokes after major base strokes	1.7	1.8	1.7	1.7	1	1
closed curve start position determines motion direction	1.0	1.0	1.0	1.0	4	1
start at leftmost possible start point	0.5	0.5	0.5	0.5	11	2
no down-to-up sequencing of base strokes	0.8	0.8	0.9	0.8	7	3
no right-to-left sequencing of base strokes	0.8	0.7	0.8	0.8	8	4
start with stroke containing point closest to upper left corner	0.2	0.2	0.2	0.2	12	5
no pen lifts	0.9	0.8	0.9	0.9	6	5
no down-to-up base strokes	0.8	0.7	0.8	0.8	8	6
no right-to-left base strokes	0.6	0.6	0.6	0.6	10	7
HEBREW PRINT						
pen up on all transition strokes	1.5	1.5	1.4	1.4	1	1

continue to adjacent base stroke without lifting the pen	1.1	1.2	1.2	1.1	4	1
curves clockwise	1.1	1.2	1.2	1.2	3	1
no down-to-up vertical base strokes	*	*	*	*	*	2
no pen lifts	1.4	1.3	1.3	1.3	2	3
no left-to-right sequencing of base strokes	0.5	0.5	0.6	0.5	8	4
no down-to-up base strokes	1.1	1.2	1.1	1.2	4	5
no right-to-left base strokes	0.4	0.3	0.3	0.3	9	6
start at leftmost possible start point	0.3	0.3	0.4	0.3	9	6
minor base strokes after major base strokes	0.9	0.9	0.9	0.9	6	7
start with stroke containing point closest to upper right corner	0.7	0.6	0.7	0.7	7	8
HEBREW SCRIPT						
pen up on all transition strokes	1.0	1.0	1.0	1.0	3	1
no high-precision intersections	1.3	1.3	1.3	1.3	1	1
first base stroke starts between baseline & x line	1.3	1.3	1.3	1.2	2	1
closed curve start position determines motion direction	1.0	1.0	1.0	1.0	3	1
no down-to-up base strokes	0.9	0.9	0.9	0.9	5	2
start with stroke containing point closest to upper right corner	0.7	0.7	0.7	0.8	6	3
start at rightmost possible start point	0.6	0.6	0.6	0.6	7	4

We ran the algorithm initializing the weights to 1 on the prescribed writing of Hebrew print and Hebrew script, and found similar results to Roman (Table 17). Once we found stable weights that yielded a successful model in HG, we could also compare them to the rankings of OT. Constraints'

HG weights clearly did not correspond exactly to their OT ranking, let alone follow the same domination relations. At a minimum, domination relations would dictate that the dominating constraint have a greater weight than the dominated one, thus posing a greater penalty on one violation of the former compared to the latter. To fully correspond to OT's strict domination, the weight of the dominating constraint would have to be greater than the weight of the dominated constraint multiplied by the possible number of violations the latter may incur.

The HG weights we have found hardly ever reflected that stricter version of correspondence to OT. Moreover, on some cases the weight of the dominating constraint was smaller than that of a constraint it had to dominate according to OT. For example, the constraint 'start at leftmost point' had to dominate 'no pen lifts' for Roman writing according to OT, but in all four of our HG runs the latter had a greater weight than the former. Nevertheless, we found great similarity between the ranking of constraints in different strata in OT and the ranking of the weights by their magnitude in HG. For Roman lower and uppercase letters, we found a correlation between the rankings in OT and HG of $r = 0.64$ ($p = 0.027$), for Hebrew print it was $r = 0.72$ ($p = 0.014$), and for Hebrew script $r = 0.93$ ($p < 0.001$). Additional work would be required to devise means of comparing OT and HG models.

The importance of individual constraints

Despite the difficulty in providing a transparent interpretation of the weights in HG, in some other areas we found our modeling results with HG less opaque than our modeling with OT. For example, one reason to use HG to supplement our OT analysis was that HG allows a fairly quick determination of the relative importance of any individual constraint to modeling a set of targets. An OT model without one of the constraints from the minimal set needed for any individual participant (or for the prescribed way of writing) would necessarily fail to find a ranking. We could still score all the possible candidates if we removed a constraint from the minimal set, but we would not be able to adjust the ranking to reflect the missing constraint and optimize the model (i.e., by giving some other constraint the chance to "shoulder the burden", at least for some letters). In HG, on the other hand, we can still learn the best weights for a subset of the constraints, and easily find out which letters fail under the best possible weights when removing one constraint. In that way we can determine not just what constraints apply to each letter (which could be achieved by looking at constraint violations), but also which constraints are individually responsible for picking a target over its alternatives.

To test the importance of individual constraints, and determine how much work each constraint does (i.e., how many letters depend on this constraint to have their target chosen), we ran the GLA for HG on the prescribed writing of English (upper and lowercase combined), Hebrew print, and Hebrew script, each time excluding one of the constraints (e.g., if there were 10 constraints in the minimal set for OT, we ran the GLA ten times, each time removing one of the 10 constraints and including only the remaining 9). We recorded the overall level of success for the subset of constraints, as well as the letters on which it failed.

In addition to the minimal sets of constraints, for which we could find weights that worked with HG, we were also able to find a set of weights that would yield all and only target stroke patterns when we removed the constraint 'no down-to-up vertical base strokes' from the Hebrew print set. This constraint was only needed for a successful ranking within the framework of OT, but was no longer needed once other constraints could team up to overcome violations of higher-ranked constraints. Interestingly, when we ran the GLA with and without this constraint on Hebrew print, we found remarkably similar weights (when initializing all the weights to 1), with one major difference: The constraint 'no down-to-up base strokes' now seemed to "pick up the slack", and got an inflated weight (Table 18).

CONSTRAINT	AVG WT WITH	RANK WITH	AVG WT W/O	RANK W/O
pen up on all transition strokes	1.4	1	1.5	1
continue to adjacent base stroke without lifting the pen	1.4	1	1.2	4
no pen lifts	1.3	3	1.3	2
curves clockwise	1.3	4	1.2	3
minor base strokes after major base strokes	1.0	5	0.9	6
no down-to-up base strokes	0.8	6	1.2	4
start with stroke containing point closest to upper right corner	0.7	7	0.7	7
no down-to-up vertical base strokes	0.7	8	*	*
start at leftmost possible start point	0.6	9	0.3	9
no left-to-right sequencing of base strokes	0.5	10	0.5	8
no right-to-left base strokes	0.3	11	0.3	9

Table 18. The minimal set of constraints needed to model Hebrew print in OT, their weight and ranking when including all 11 constraints (averaged across four runs of the GLA), and the average weight and ranking when excluding the constraint 'no down-to-up vertical base strokes', and including only the 10 constraints needed for a successful model in HG. Avg = average; wt = weight; w/o = without the 11th constraint.

Since we found out that the constraint 'no down to up vertical base strokes' was not necessary for the modeling of Hebrew print, we removed it from the set and ran the script again, excluding each of the remaining 10 constraints in turn. We were subsequently not able to find a smaller set of constraints that would account for all of the data, in any of the three scripts. Our success rate when removing one constraint ranged from 81-98% (mean 88%) for Roman characters, from 48-97% (mean 85%) for Hebrew print, and 47-93% (mean 78%) for Hebrew script (Table 19).

While removing each of the constraints in English had a relatively low impact, causing between 1 and 10 targets (2-19%) not to be chosen correctly, in Hebrew there was one constraint that, when removed, caused an avalanche of wrongly-selected non-targets to be chosen as the most Harmonic. In Hebrew script, six of the seven constraints had an impact on 2-7 letters (6-21%). The last constraint, 'no down-to-up strokes', caused a failure of 16 letters (53%) when it was excluded from the modeling. Similarly, in Hebrew print, nine of the ten constraints needed for a successful model with HG caused a failure in 1-8 letters (3-24%) when removed, but again removing the constraint 'no down-to-up strokes' caused a failure on 17 letters (52%), significantly more than expected by chance (χ^2 = 54.2, p < 0.001; R_i = 4.11, p < 0.001).

Table 19. The constraints needed for a successful modeling of the prescribed writing of Roman lower and uppercase (top), Hebrew print (middle), and Hebrew script (bottom). For each constraint we indicate the maximum accuracy (as a percentage of letters whose targets are chosen) reached when it was excluded, and the letters on which the modeling fails without it. The number in parentheses indicates more than one allograph of a certain letter failed. Max acc = maximum accuracy; N fail = number of letters whose targets were not chosen.

CONSTRAINT	**MAX ACC**	**FAILED LETTERS**	**N FAIL**
ROMAN UPPER AND LOWERCASE			
pen up on all transition strokes	88%	G, N, V, W, v, w	6
no initial down-to-up base strokes	96%	A, N	2
minor base strokes after major base strokes	88%	A, E, F, H, I, Y	6
no down-to-up curved base strokes	98%	D	1
closed curve start position determines motion direction	83%	O, Q, a, b, d, g, o, p, q	9

start at leftmost possible start point	98%	Y	1
no down-to-up sequencing of base strokes (v2)	92%	E, F, I, Y	4
no right-to-left sequencing of base strokes (v1)	87%	B, D, M, N, P, d, m	7
no pen lifts	88%	G, N, V, W, v, w	6
start with stroke containing point closest to upper left corner	98%	g	1
no down-to-up base strokes	85%	H, X, Y, a, d, q, x, y	8
no right-to-left base strokes	83%	A, E, F, H, I, T, f, r, t	9
HEBREW PRINT			
pen up on all transition strokes	76%	ב, ל, נ, ע, פ, ש, ת, ף	8
continue to adjacent base stroke without lifting the pen	97%	ת	1
curves clockwise	94%	ס, ש	2
no pen lifts	100%	ב, ל, נ, ע, פ, ש, ת, ף	8
no left-to-right sequencing of base strokes	76%	א, ע	2
no down-to-up base strokes	94%	א, ב, ג (2), ד, ה, ו, ז, ח, ל, ע, צ, ק, ש (2), ת, ץ	17
no right-to-left base strokes	48%	ב, ת	2
start at leftmost possible start point	94%	ז, ס, ם	3
minor base strokes after major base strokes	91%	צ	1
start with stroke containing point closest to upper right corner	97%	ב (2), ז, ש (2), ם	6
HEBREW SCRIPT			
pen up on all transition strokes	80%	א, ה, ח, ק, ת, ם	6
no high-precision intersections	77%	ב, ז, ל, מ, פ, ף, ץ	7

first base stroke starts between baseline & x line (v1)	87%	ט, ל, ף, ץ	4
closed curve start position determines motion direction	87%	ס (2), ם (2)	4
no down-to-up base strokes	47%	א, ג, ד, ה, ו, ז, י, ב, מ, נ, ע, פ, ק, ש, ך, ם	16
start with stroke containing point closest to upper right corner	77%	ד, ה, ח, פ, ק, ת, ם	7
start at rightmost possible start point	93%	ע, ש	2

Some constraints exhibited similar patterns across the three different scripts, indicating that their applicability was not limited to a particular set of shapes. For example, removing the constraint 'pen up on all transition strokes' caused a similar level of damage in Roman, Hebrew print, and Hebrew script letters (12%, 24%, and 20% of letters failed without it, respectively). Most constraints' removal caused a failure in 6-24% of characters. However, other constraints affected significantly fewer or more letters than that. For example, 'no down-to-up curved strokes' in English was necessary only for one character: D. Similarly, the constraint 'minor strokes after major strokes' was only needed to account for the letter Tsadi (צ) in Hebrew print. The relatively limited effect of these constraints could stem from an uncommon letter shape, but it could also mean that the constraint in question is too specific.

At the other end of the scale are constraints that cause a disproportional amount of damage. As we mentioned earlier, 'no down-to-up strokes' caused a failure of over 50% of letters in Hebrew print and Hebrew script when it was removed. Significantly, the constraint dictating no down-to-up movement can also be considered the most general, in that it applies not just to the writing of letters, or even to the drawing of non-letter shapes, but also to well-controlled hand movements in general, regardless of the hand used. Excluding this constraint admittedly caused a failure in only 15% of letters in the Roman set, but the minimal constraint set for Roman letters also included the constraints ‘no down-to-up curved strokes’ and ‘no initial down-to-up strokes,’ which may have prevented the model to fail more severely when ‘no down-to-up strokes’ alone was excluded.

Participant writing

Another reason to use HG was our hope that we would be able to account for the handwriting of participants whose stroke patterns we could not model using OT. We attempted to model the handwriting of the 40 participants described in Chapter 3, and searched for a set of weights that would yield the greatest Harmony within the framework of HG for all and only target stroke patterns compared to the non-target alternatives for each participant. As expected, we managed to find such a set of weights for the 20 participants for whom we found a ranking of constraints with OT (10 RHs, 10 LHs). This was not surprising, as OT rankings can, in a finite tableau with a known number of constraint violations, be directly translated to a set of weights that would work with HG. In addition to the 20 participants we could already model with OT, using HG we found a set of weights that chooses all and only target stroke patterns for an additional four participants (2 LHs and 2 RHs). For two of those participants we had partial success with OT (one LH participant who we could model without including the letter y, and one RH participant we could model without x). With the remaining two participants we could not find a ranking with OT at all.

When we allowed negative weights, our average success rate was 97.5%, or an average of 0.57 failed letters per participant (with full success for 24 of 40 participants, as described above). When we ran the GLA again, limiting the weights to non-negative values only, we achieved an average success rate of 97.2%, or 0.62 failed letters per participant (with full success for 22 of 40 participants). In looking at just the number of participants we were able to fully model, switching from OT to HG allowed us to account for two additional participants, and allowing negative weights helped us account for another two. The two participants we could only model when allowing negative weights each had a strong negative weight associated with at least one of their constraints.

Participant AMD, a LH female whose handwriting we were able to model in OT when removing the letter y, needed both ‘no down-to-up strokes’ and ‘no right-to-left sequencing of strokes’ to have a negative weight. Those reversed constraints were needed to account for three letters she wrote differently than most participants: f, which she started from the bottom, x, which she started from the top right and then produced the second stroke from the bottom up, and y, in which she produced the longer stroke (on the right) first, and the shorter one second. Participant BKD, a LH male, needed a similar combination of constraints with negative weights: ‘no down-to-up sequencing of stroke start positions’, and ‘no right-to-left sequencing of strokes.’ Interestingly, participant BKD also produced his x and y the same way as AMD. The negative weights given to the constraints described above do not exactly correspond to any of the constraints we have formulated in our set, and so could not yield a full ranking in OT, or in HG when weights were limited to non-negative values.

Modeling the field

Using HG, we also attempted to model the entire field of stroke patterns that participants produced. RH and LH participants produced a total of 80 different stroke patterns for the 41 Roman lowercase allographs. For the letters o and c, participants produced all possible stroke patterns (e.g., for o, 35 participants produced it counter-clockwise, and 5 participants produced it clockwise). For other letters, such as w, and the z allograph with the horizontal crossing line in the middle, participants produced only 0.04% or 0.07% of possible stroke patterns, respectively (e.g., every single participant produced lowercase w the same way, despite there being 2542 possible stroke patterns). Overall, the 80 stroke patterns participants produced represent only 1.26% of the total number of possible candidates for Roman lowercase letters (6360 candidates).

We attempted to find a set of weights that would account for all and only the target stroke patterns ever produced by our participants using HG. We ran the algorithm three times: Once on the combined set of all 80 targets produced by RH and LH participants, once on the 53 targets produced by RH participants only (0.84% of all possible candidate stroke patterns), and once on the set of 73 targets produced by LH participants (1.15% of all possible candidates). We were only able to achieve a maximum accuracy of 65% on the combined set of targets. Because some participants prefer to produce horizontal strokes from left-to-right and others from right-to-left, for example, it makes perfect sense that we would not be able to model them together.

When we modeled RHs' and LHs' writing separately, we were able achieve a maximum accuracy of 67% on the set of targets for LHs (even when excluding those participants who produce horizontal strokes from left-to-right), compared with a high of 92% of letters (all but k and t) for RHs' targets. The high level of accuracy for RHs combined with the low level of accuracy for LHs is perhaps indicative that RHs' writing is more consistent across participants than LHs'. The reasons for RHs' writing being more homogeneous across participants can probably be traced back to their instruction. Whereas LHs were almost certainly taught to write in a way that was not optimal for their preferred effector, and thus had to make adjustments to their writing, RHs did not face such additional obstacles in developing their writing style. Indeed, RHs' stroke patterns differed from the prescribed by an average of 11.6 letters (SD 1.3), and LHs' stroke patterns differed from the prescribed by an average of 14.5 (SD 1.9). This difference was statistically significant ($t(38) = 5.35$, $p < 0.001$).

Odd-ball letters

When attempting to model participants' handwriting using strict domination (i.e., in OT), if a ranking did not exist our model simply failed. Using HG, we were able to edge as close as possible to success, improving the model iteration by iteration even if it could not account for all of the data. We were therefore able to gain insight into which letters' stroke patterns are less consistent with the rest for a given participant. In theory, there could be a situation in which two letters' stroke patterns stand in direct opposition to one another but both are compatible with the rest of the alphabet. Thus, it is possible that two models will achieve the same level of success, each of them failing on just one letter, but that letter be different between the models.

While there certainly were cases in our data in which the same rate of success could be achieved without either of two letters, the reality was that we saw the same letters failing over and over again in different participants, despite those participants using different stroke patterns, and the algorithm being initialized with a different seed for the random sampling of letters on which to update the weights. When looking at the most generous HG model (allowing negative weights), we could not find a set of weights that would choose all and only target stroke patterns for 16 participants. Of them, 10 participants failed on the letter k, and 6 failed on the letter y. We are unsure about the reason why some letters are more prone to failure than others, but it is worth noting that the letters k and y in particular saw great variability in the stroke patterns used to produce them. Participants used a total of 8 different stroke patterns to write k, and 5 different stroke patterns to write y. And while the most common stroke pattern for each matched the prescribed stroke pattern, and was rarely the one that caused failure, the failures were also distributed across many stroke patterns.

Both the failure to model and the variability in stroke patterns may stem from those letters having properties atypical of Roman letters. Both k and y have two features that are rare in other lowercase letters: They both have diagonal strokes that meet (but do not cross) at a non-right-angle (as do v and w, although there the intersection is between the ends of two lines, whereas in y and k the intersection is between the end of one line and the middle of another). We know from our modeling that people prefer to avoid intersections such as those in k and y (as is evident by participants' use of the constraint 'no high-precision intersections'). And while participants cannot avoid diagonal lines if those appear in the base-shape, those lines might present particular challenges as they include both a change of position on the x-axis and on the y-axis. For example, diagonal lines can create a conflict in preferences if, say, a top-down movement forces the writer to produce a right-to-left movement (as in the top-right stroke in k and the rightmost stroke in y). In all other Roman lowercase letters where a conflict between the up-down axis of movement and the left-right axis of movement emerges, the conflict can be resolved by preferring the direction in which there is no pen lift (e.g., continuing in an upwards direction on the second stroke of v), but not on k and y.

Discussion

In this chapter, we described a HG analysis of the data of RH and LH participants writing English using their dominant hand, and of the prescribed writing in English (lower and uppercase), Hebrew print, and Hebrew script. We used HG because it allows us to answer more fully some of the cognitive questions we posed earlier. Our occasional inability to model some participants' handwriting with OT raises broad questions about the overall success, or lack thereof, of the approach. As we mentioned earlier, one of the options is that these failures to model are due to inherent inconsistencies in participants' handwriting. Participants may not use rules at all to guide their production of stroke patterns in writing, or their set of constraints might not apply in a consistent way across all the characters in a language, or perhaps even across multiple productions of the same character. Alternatively, and in our opinion more plausibly, our occasional failures to model could stem not from some inherent flaw in the application of the framework as a whole, but rather from some feature of our model. We have discussed the possibility of individual constraints being formulated in a way that is incompatible with their representation in participants' minds. But in this chapter, we wanted to explore the possibility that (some or all) of our failures to model stem from the nature of strict domination, which might not fully capture how people produce stroke patterns.

HG is very similar to OT, as it relies on the same kind of violable constraints, but it does not require constraints to be ranked in strict domination. Instead, HG assigns each constraint a weight, and calculates candidates' Harmony based on the weighted sum of constraint violations. Because we were able to use the same constraints with our HG model as with our OT model, the comparison of the models' performance is relatively straight-forward, and any differences between them can be traced back to the requirement of strict domination. We applied HG to attempt to model the writing of the 40 Participants whose handwriting we modeled using OT in Chapter 3. Using OT, we managed to find a ranking of constraints that would account for all and only target stroke patterns for 20 of those 40 participants (10 LHs and 10 RHs), and we managed to find a ranking that accounts for all but one target for an additional 10 (4 LHs and 6 RHs). The fact that our success was not complete with all of the participants allowed us to test our OT model against the HG model.

As expected, we were able to find a set of weights that would account for the writing of the 20 participants we were previously able to model using OT. However, we were also able to find a set of weights that would account for the writing of an additional 2-4 participants (depending on whether we allowed non-negative weights), for whom we failed to find a ranking in OT. The ability to model additional participants beyond those we were able to model with OT suggests that the concept of strict domination was holding us back, at least in some cases. However, because we were not able to answer some of the questions that interested us using HG along, we argue that HG is not necessarily preferable to OT for the purpose of exploring the cognitive mechanisms underlying handwriting.

OT vs HG

The goal of this work was not just to model as much of the data as possible within a certain framework, but to use the modeling to gain insight into the cognitive mechanisms underlying handwriting. More than showing that “a ranking exists” that accounts for the data, we wanted to interpret the ranking or the use of constraints in terms of the causal structure governing the stroke patterns used in writing. Interpreting the data is challenging both in OT and in HG, but the challenges manifest differently in each framework.

With OT, it is tempting to think that a ranking of constraints is a straight-forward result that can be easily interpreted: Constraints that are ranked higher are more important than those ranked lower. But this is actually not always the case. As we explained earlier, our ranking algorithm is in fact a *constraint demotion* algorithm, that only demotes a constraint to a lower rank if it has to be dominated by another. Thus, constraints that are ranked higher could actually be largely irrelevant to the set of targets, or they could be required to dominate many or all other constraints. Determining what has to dominate what is a tough problem, that is not immediately solved simply by looking at the ranking. Furthermore, constraints that do not make it into the “minimal set” that we have defined, meaning that we managed to achieve a successful ranking without them, might still have to be dominated by other constraints for the modeling to work. HG presents different challenges. In the way that we implemented the algorithm (a GLA that searches for the optimal weights), we were not guaranteed to find a vector of weights that would yield all and only target stroke patterns, even if such a vector did exist. Even once a vector of weights is found, the task of interpreting the weights is not straightforward. A constraint might be weighted heavily because it is more “important”, or because there is a small difference in the number of violations between the target and some non-target alternative.

In the end, we have found both frameworks to be useful for understanding the cognitive mechanisms underlying handwriting, each with its own advantages and disadvantages. OT allowed us to learn about the principles involved in deciding which stroke pattern to produce, and how conflicts between those principles are resolved. It also allowed us a relatively quick method of determining the smallest set of constraints that would yield a successful model (compared to HG, which has to learn new weights with each new combination of constraints). HG gave us insight into the importance of each constraint, and how broadly it applies to our data. It also illuminated some of the limitations of our implementation of OT, and cases where the use of a more lenient weighting mechanism (i.e., without imposing strict domination), could result in a slightly more compact model, with fewer constraints.

General Discussion

The production of letter-strokes in handwriting can vary widely – every letter-shape can be produced in many different ways – and yet we see a great deal of consistency, and signs of organizing principles in the way people write letters. In this dissertation, we attempted to model those principles, and to learn about the underlying mechanisms governing handwritten production. We applied OT (Prince & Smolensky, 1993) and HG (Legendre et al., 1990a), and asked whether a framework of violable constraints, that are either ranked in strict domination, or else weighted, is useful in explaining the way letters are written. We first looked at the prescribed way of writing both English and Hebrew and modeled it using OT. We then asked whether constraints ranked in strict domination can be used to model not just the way people are taught to write, but the stroke patterns they actually produce. We examined the writing of adult Hebrew and English-speakers, and of right- and left-handed participants. Finally, we used HG to examine both prescribed and participant writing, and to draw further conclusions about how the constraints are applied.

One of the main questions of this work was whether there is sufficient systematicity in the stroke patterns people use to produce characters that could be explained using ranked, violable constraints. Whether or not we can model handwriting of letters within the framework of OT is an empirical question that gets answered in this study. However, we strived to take the analysis several steps forward, and asked what we can learn from the model and the constraints involved in it. We wanted to understand whether we can use the constraints and their rankings to learn about the different principles involved in writing in different languages or by different-handed people, about how conflicts are resolved when two principles dictate contradicting stroke patterns, and whether we can characterize differences in stroke patterns (e.g., between different people) using a different ranking of constraints.

We first examined the differences in the prescribed writing of English, Hebrew print, and Hebrew script using the ranking of constraints. Due to the nature of the constraint demotion algorithm, the ranking of constraints is not always directly tied to their necessary domination relations. A constraint only gets demoted to a lower rank if another constraint has to dominate the first to pick a target over a non-target. Therefore, constraints ranked in the highest stratum may not have to dominate any other constraints at all. Furthermore, a constraint ranked in the lowest stratum might only have to be dominated by a constraint in the immediately higher stratum, and not necessarily by any other constraints.

We therefore utilized the Fusional Reduction algorithm (FRed; Brasoveanu & Prince, 2005) to determine what necessary domination relations exist between constraints. Using FRed, we were able to compare the hierarchy of the constraints, and by extension, the hierarchy of the principles involved in English and Hebrew writing. We were able to shed light on some previously unanswered questions, such as how the conflict between overall reading and writing direction and the direction of best-controlled movement is resolved in Hebrew. Contrary to what has been suggested by researchers in the past (e.g., Van Sommers, 1984), Hebrew letters are neither written entirely from left-to-right (in accordance with the direction of best-controlled movement) nor entirely from right-to-left (in accordance with the direction of transition between letters). Instead, both 'no right-to-left strokes' and 'no left-to-right sequencing of strokes' are needed to account for the way Hebrew is written (but the latter must dominate the former).

We then examined the writing of Hebrew and English-speakers writing their respective native language. We found, again, that we can model the writing of a large majority of them using ranked violable constraints. By comparing domination relations, we were able to explain why participants' stroke patterns differ. We highlighted this by directly contrasting participants whose handwriting differed only very slightly (by one letter's stroke pattern), and showing that their constraints' domination relations similarly differed only by a little. We then showed that even when there was a huge difference between the stroke patterns produced, for example between a participant's writing and the prescribed stroke patterns (differing in our data by as many as 16 letters' stroke patterns), we sometimes needed as few as two opposing domination relations to account for the difference. We further used the results of this analysis to learn, from the necessary domination relations, which principles are more important in actual vs. prescribed writing, or among different participants.

Finally, we looked at the writing of right-handed (RH) and left-handed (LH) participants writing in English. We achieved a similar level of success in modeling these participants' handwriting as we did with the English-speakers from the previous analysis. We also achieved a similar level of success in modeling the writing of RH and LH participants. By modeling RH and LH participants, we were again able to examine what happens when different principles are dictating different stroke patterns. As in Hebrew, in which there was a conflict between the overall reading and writing direction and the direction of best-controlled movement, in LH English-speakers the same conflict arose, only in the opposite direction. For LH English-speakers, the direction of best-controlled motion for horizontal strokes is right-to-left, whereas the direction of transition between letters in English is from left-to-right. And again, similarly to Hebrew, here too we have found that LH English-speakers needed both 'no left-to-right strokes' and 'no right-to-left sequencing of strokes' to model their handwriting.

Our analysis also served to highlight the fact that for RH English-speakers neither 'no right-to-left strokes' nor 'no right-to-left sequencing of strokes' was sufficient for a successful model. Like the LHs and the prescribed Hebrew print, RHs too needed both a constraint governing the direction of strokes, and one governing the sequence of strokes. Our modeling here is particularly illuminating seeing that in RH English-speakers these two constraints were compatible, both dictating a left-to-right direction, and so without our modeling this point might have been missed.

Our modeling of handwriting with OT marks a significant advance over the research that has been done thus far. While several researchers before us have described the organizing principles governing handwriting (e.g., Goodnow & Levine, 1973; Goodnow et al., 1973; Nihei, 1983; Van Sommers, 1984), none of them were able to deal fully with the issue of conflicting principles or the fact that the rules derived from the principles are often violated. By shifting the discussion from a discussion of hard-set rules to one of violable constraints, we took the first step to dealing with conflicts and violations. We were able to show not just that the conflicts between constraints *can* be resolved, but also *how* they are resolved – in a systematic and non-arbitrary way that is consistent across letters.

We have found OT immensely useful in explaining the systematicity and variability of stroke patterns in handwriting. We have further found it an excellent tool to enhance our knowledge and understanding of the principles underlying handwriting, and the various ways those principles interact. However, we were not always able to model participants' handwriting within the framework of OT. In the next section we discuss several reasons why our modeling efforts were not always successful, including reasons related to our definition of constraints and targets. We also compare the benefits and disadvantages of OT and HG for the modeling of handwriting. We ask whether relaxing only the requirement of strict domination, which is the main difference between OT and HG, might create a more robust model of people's handwriting.

Limitations of our model

The modeling, although successful to a large extent, did not succeed in every case. In this section we examine the reasons why it occasionally may fail. First, for our handwriting data to be modellable within OT, certain requirements have to be met. 1) There needs to be some systematicity in stroke patterns with respect to the governing principles used (i.e., the same governing principles should apply in a consistent way across different characters within a language). 2) We need to have a reasonably good characterization of what those governing principles are, and how they apply to the stroke patterns. And 3) for our particular model to work (i.e., for us to be able to model stroke patterns in writing specifically within the framework of OT), the principles need to have a certain nature: They need to be *violable constraints*, and, should they come into conflict, this conflict should be solved by *strict domination*. For our model to be scientifically sound, those violable constraints should arise from general principles that are broadly applicable and externally motivated (i.e., are not defined ad-hoc for a particular character).

When we fail to find a ranking of constraints that would account for a participant's handwriting, one may assume that one of the three requirements above was not met. However, to have a successful model, in addition to the data's internal structure, we also need correct representation of the target stroke patterns, and a correct definition of each of the constraints. We took it as a working assumption that constraints, phrased in our implementation in invariant terms (e.g., "left-right", and "up-down"), come into play when developing the effector-independent motor plans. However, the correct formulation of each of the constraints (i.e., the exact logic which is applied in each constraint) did not stem directly from the aforementioned assumptions regarding constraints being violable and ranked.

In addition, even if we assume that our representation of the base shape of letters is correct, and that we derived all the possible stroke patterns from this shape appropriately, it is not guaranteed that we will have correctly identified the target stroke patterns participants used. Only if we are able to correctly characterize both the general structure of the system underlying handwriting (as described above), the target stroke patterns used, and the constraints that apply to these stroke patterns, can we expect our implementation of the model to succeed. If not, we expect it to fail. Below, we discuss in greater detail our challenges in identifying targets and formulating constraints.

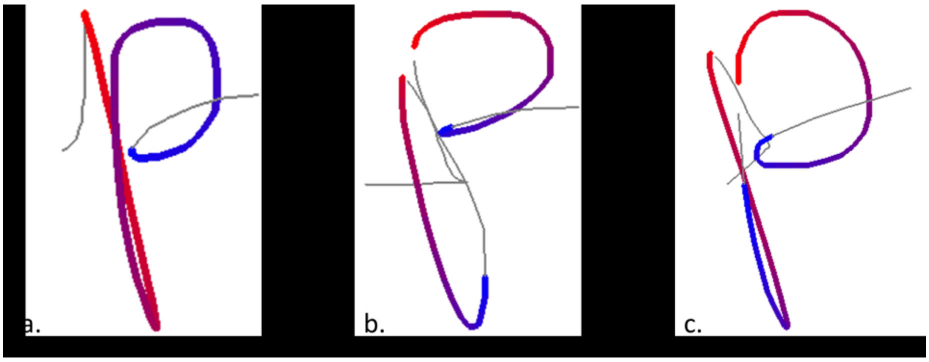
Limitations of our selection of targets

One of the reasons why we might fail to model a participant's handwriting is that we have incorrectly identified the targets they are using. Three aspects of the definition of targets should be considered: 1) Whether we have identified *all* target stroke patterns that a participant uses; 2) Whether we have identified *only* target stroke patterns that a participant uses; and 3) Whether we have correctly identified the mental representation corresponding to a letter-shape and its constituting strokes.

The main concern in identifying *all* target stroke patterns is that our data set might be too small or limited to expose the full field of stroke patterns used by each participant. Participants may use multiple stroke patterns for some letters, and we may or may not observe all of them within one session of our experiment. For example, someone might have two stroke-pattern representations for the uppercase letter 'A' – one starting at the top, and the other starting at the bottom left. They may produce the first stroke pattern (starting at the top) more frequently, but still produce the second one (starting at the bottom left) on some occasions. For example, they might begin the letter 'A' at the bottom after a previous letter ended at the bottom (e.g., after writing Z), or at the beginning of a sentence. They might even not have any particular reason to produce one stroke pattern over the other, and we happen to only see one of them within our data collection session by chance. The possibility of having multiple stroke patterns for a particular letter becomes more of a problem when considering that some of the less frequent letters in the alphabet were only sampled in our word list 4 or 5 times.

In terms of correctly identifying *only* target stroke patterns, the problem is distinguishing performance issues from cases of actual different representations. For example, it is pretty clear that the lowercase 'p' in Figure 16a was written without lifting the pen, and that the lowercase 'p' in Figure 16b was written with a pen lift between the two strokes. But what about the character in Figure 16c? It is unclear whether the participant intended to lift the pen, but did so lazily, lifting the pen only a little bit, and creating a mostly-pen-down transition stroke, or whether they intended not to lift the pen, but produced a weaker stroke, and so created a gap between the end of the transition stroke and the beginning of the curved base stroke.

Figure 16. Three instances of the lowercase letter 'p' written by the same participant on different trials. The letters were produced in the words temper (a), grape (b), and nephew (c). The color changes from red, marking the beginning of the stroke, to blue, marking the end of the stroke. The thin gray lines indicate a pen-up movement.



Correctly identifying all and only target stroke patterns for a given participant is crucial, because our ranking algorithm requires that all the targets be more Harmonic than all non-targets, but it does not specify a ranking among targets (i.e., if a participant produces two different stroke patterns we only require that they both be chosen over all others, but we do not require them to have the same Harmony, or for any specific one of them to win over the other). Therefore, it is quite possible that a ranking exists that chooses two stroke patterns over all others, say the pen-down stroke pattern for lowercase p which appears in Figure 16a, and the pen-up stroke pattern for the same letter (Figure 16b). However, if one of the two stroke patterns is removed from the set of targets, suddenly the ranking fails, because the remaining target is no longer the sole winner.

Even if we have identified correctly all and only the stroke patterns for a given participant, there is a chance that the algorithm would fail, and not because the framework is incorrect or because the constraints are not formulated properly. Instead, those inconsistencies might stem from differences in constraint rankings based on some external factors. For example, a participant who feels rushed might value speed in her stroke patterns more than she would if she was trying to write more legibly, and thus, when in a hurry, she might produce more pen-down transition strokes than she would when taking her time. Anecdotally, we have noticed a tendency of participants to keep the pen down more on transition strokes towards the end of our hour-long experimental session. To test the changes in stroke patterns more thoroughly, and to find out whether we can separate them into subsets which we can model within the OT framework, we ran an experiment in which we directly manipulated participants' preference by instructing them to write “more quickly” or “more neatly”.

The results of this experiment are outside the scope of this dissertation, but will be reported in a future paper.

Importantly, we note that the possibility of small changes in constraint rankings does not necessarily stand in contradiction to our assumption that constraints are activated only when learning the motor plans (or stroke patterns), rather than online every time one writes. First, learning in this context does not have to be limited to a critical period, but can span as long as the person's writing style is developing. But also, a reasonable possibility is that participants may have developed two sets of motor plans, based on slightly different rankings of constraints, that they can retrieve in the appropriate context. Our experiment testing "neat" vs. "fast" writing could illuminate this further.

Finally, there may be considerations not directly related to the principles underlying handwriting for how people write certain letters. For example, a letter that appears with a lower frequency may manifest a less stable stroke pattern, with people "trying out" different stroke patterns on different occurrences (as seems to be the case with the word-final form of the letter Tsadi in Hebrew). Another external influence might come from a person's admiration for how their best friend, or their idol, wrote some letter in grade school. And the presence of a certain letter in someone's name might affect how it is written in their signature, and eventually change how they write it in general. In particular, all those external circumstances might affect the writing of one, or a few letters, but not others. While we do not want to give ourselves too many degrees of freedom, such that we can dismiss any results that do not fit with our interpretation as mere noise, it is nevertheless not unreasonable to assume that our data, and any handwriting data for that matter, will contain a certain amount of irreducible and unexplainable noise, that nevertheless does not render the framework unviable.

Limitations of our constraints

A failure to model might also stem from incorrect characterization or formulation of the constraints. Even if participants do follow basic principles when writing, and if those principles are ranked, violable constraints, if the constraints are not phrased correctly, we would not be able to capture their effect. Without proper definitions of the constraints we might fail to find a ranking that chooses all and only target stroke patterns, despite there being one. Many constraints can be defined in different ways that tap the same basic principle but manifest slightly differently. For example, the constraint 'no right-to-left sequencing of strokes', which concerns the order in which strokes are produced, could refer to the end of the first stroke and the beginning of the second, or to the starting positions of the two strokes, or their ending positions. Moreover, that constraint can also refer to two sequentially adjacent strokes, or to any pair of strokes within a letter.

The many decisions that have to be made regarding the formulation of constraints do not reflect only the problem of finding the "correct" representation of the constraint, but also the fact that different people might have slightly different manifestations of the same basic constraint. Despite following the same principle, different people might apply it slightly differently (e.g., 'no down-to-up strokes' may refer to any stroke with a vertical component, or only to completely vertical strokes). While there could be one universally successful definition of the constraint, it is also possible that different instantiations of the constraint are necessary to account for the handwriting of different people, or in different languages. When we encounter a participant whose handwriting we cannot model, we have to wonder whether their handwriting is not internally consistent, or whether we do not have the correct constraints to characterize the consistency in their handwriting. And the missing constraints could be ones we have not thought of at all, or they could just be slight variations of constraints we do have.

Of course, another possible explanation for why our success in modeling participants' handwriting was not complete is that the set of constraints we used is not entirely comprehensive. One way to combat this occasional failure to model a participant's handwriting would be to add additional constraints. Successful modeling of handwriting within the framework of violable constraints could theoretically be based on very targeted constraints, such as "when encountering the letter 'T' start it at the top and produce the vertical stroke first, and then the horizontal stroke from left-to-right."

Such a constraint would very obviously not be generally motivated or broadly applicable across characters, but some less obvious instances of constraints that are too targeted could unintentionally make their way into our set of constraints. We took steps to ensure that this is not the case by providing broad external motivation for each of our constraints, and by verifying that constraints are applicable to as large a set of characters as possible, at the cost that we may have omitted some constraints that would be necessary for the modeling of some participants' writing.

One area where our constraints might be under-representing the mechanisms that people use is in looking at the surrounding letters. All of our constraints apply exclusively to the strokes within a single letter, and none of them considers where a previous letter ended or the next one starts. Since we know (e.g., from graphemic buffer dysgraphia, Hillis & Caramazza, 1989) that participants hold more than one letter in mind when writing, it is reasonable to assume that the stroke patterns would be affected by the surrounding letters. Furthermore, there could be effects related to a letter's position within a word that are themselves independent of the previous or next letter (e.g., different stroke patterns for word-initial than for subsequent letters). Further research is warranted to determine if adding such considerations significantly improves the rate of success of our model.

The process by which we chose the constraints for our implementation of OT included some trial and error (adding constraints or adjusting their exact definition to account for occasional failures to model), but was ultimately motivated by as general and broadly applicable principles as possible. While this process is not complete, we felt that the level of success we achieved in modeling participants' handwritten stroke patterns is well-balanced with the fact that our individual constraints are applicable to many characters, and well-motivated by general principles.

Is the framework viable

In order to discuss our rate of success in modeling participants' handwriting using OT, we had to confront the question of what constitutes "success". Many computational models of cognitive function are measured against the level of success achieved by previous modeling attempts (e.g., Lake et al., 2012, 2015). This does not seem to be possible for our endeavor for two reasons: First, to the best of our knowledge, there have not been any attempts at comprehensively modeling the stroke patterns of individuals across an entire set of characters, let alone while comparing different languages or different dominant hands. Secondly, and perhaps more importantly, our success in modeling an individual's handwriting using OT is "all or none". We either find a ranking of constraints that accounts for all and only target stroke patterns for that individual or we do not.

While in some cases we have come closer than others (e.g., achieving success when removing only one letter from the set of targets; what we have termed "partial success"), using OT we were nevertheless only able to specify "success" or "failure". We would naturally not want to characterize our entire enterprise as a failure if we failed to find a full ranking of constraints that accounts for the handwriting of a few of our participants. To address this issue, as well as the possibility that our failure to model some participants' handwriting stemmed from our requirement that constraints be ranked in strict domination, we implemented a version of Harmonic Grammar (Legendre, Miyata, & Smolensky, 1990). HG deals with complex relationships between constraints, like OT, but in HG the strict domination requirement is replaced with numeric weights representing the "strength" of each constraint. In this way, multiple violations of lower-ranked constraints could in fact be worse than a single violation of a higher-ranked constraint, giving some more flexibility to the modeling.

HG provided a benchmark against which to test our level of success with OT, as well as a way to easily learn things about the data that would require significant effort to do with OT. Using HG, we were able to quantify the level of success of the model for each individual participant, even if we were not able to account for all of their stroke patterns. Where our OT implementation would fail to find a ranking, HG allowed us to continue adjusting the weights until we could model the greatest number of stroke patterns for a given participant. We were able to determine, for example, that a certain set of weights would account for 90% of a particular participant's stroke patterns. Of course, this would have been possible with OT as well, by removing one target in turn until we found a ranking, or if the ranking failed with 25 of 26 letters, continue to removing 2 targets each time. But that would require significant computational time, and in fact might not be feasible on some instances where the success rate is relatively low.

Secondly, with HG we could fairly straightforwardly test for how many letters each constraint is crucial, by removing one constraint in turn, and testing how many targets fail with the new subset of constraints. Again, we could have implemented a version of this in OT, by removing one constraint from the ranked set of constraints, and scoring each of the targets using the new set. However, a change in constraint rankings could plausibly yield a better result after removing one constraint, and a new ranking is not possible to find in OT if the excluded constraint was part of the minimal set. Lastly, using HG we were able to model the handwriting of a handful of participants whose stroke patterns we were not able to model with OT. Furthermore, we actually found a set of weights that would explain the prescribed writing of Hebrew print using one constraint fewer than the minimal set we have found with OT. While our success with HG where we failed with OT might indicate that HG is a better model for the way people write, some of the advantages of OT over HG suggest that OT is still a more useful framework for some aspects of investigation.

Unlike with HG, in OT we were able to examine necessary domination relations. As we explained previously (in Chapter 4), the weights in HG might vary for many reasons, and a higher weight does not necessarily indicate a more important constraint. While the slightly opaque relation between rank and importance stands in OT as well, in OT we could directly find if one constraint must dominate another to achieve a successful ranking. If such a domination relation exists, the dominating constraint is, in a sense, more important than the dominated one. With the necessary domination relations, we could compare not just the ranking of constraints for two individuals, but also the ranking across two languages. In HG we would expect the weights to differ when various shapes are involved whose number of constraint violations might differ significantly, and similarly we might observe a difference in the ranking of the constraints, but the necessary domination relations in OT will stand regardless of the number of violations or even whether a constraint is required for a ranking or not.

Furthermore, we found the constraint demotion algorithm slightly more robust than the GLA for our purposes, as we were guaranteed to find a solution if one existed, whereas our GLA in HG could get caught in local minima, depending on the initialization of some of our variables (e.g., the initialization of the weights). We therefore recommend that future research aimed at understanding the principles at work during writing, as well as their violations and conflicts, utilize OT to find necessary domination relations, and possibly supplement the modeling with HG, to handle cases in which a ranking with OT is not possible.

Are some letters more susceptible to failure?

Some letters were much more prone to failures of our modeling than others. The letters y (both upper and lowercase) and k in English, and ץ (the word-final form of Tsadi), and ט (Tet) in Hebrew print, were particularly challenging to model, both in the prescribed writing and in participants' writing. Interestingly, these were the letters that participants most often produced in multiple ways on different occasions (y, k) or completely forgot their shapes (final-Tsadi, Tet). All four letters appear relatively infrequently in their respective languages (y accounts for 1.5% of letters in English, k accounts for 0.6%, Tet accounts for 1.5% in Hebrew, and final-Tsadi for about 0.1%, according to simia.net's count of letters in Wikipedia in different languages, Vrandečić, 2012). Since Hebrew speakers very rarely produce Hebrew print in handwriting, they are exposed to the writing pattern of final-Tsadi and Tet even less often than that, if at all after 1st grade.

For final-Tsadi in particular, participants occasionally did not seem to have an explicit representation of the letter at all, reacting in stumped silence to the instruction to produce it in Hebrew print, and occasionally producing the shape in Hebrew script, or producing the non-final form of the letter instead. We have previously found (Wong et al., 2018) that participants who had no trouble reading the looptail allograph of the letter 'g' (the allograph found in Times New Roman font, for example), were nevertheless often unaware that this allograph existed. They were almost uniformly unable to produce it correctly, even after being told it existed, or after actively searching for it within a text. Similarly, participants had no trouble recognizing the Hebrew print word-final form of the letter Tsadi, but they had in our data set a particularly hard time producing it correctly when prompted.

The lack of explicit knowledge about a letter, and the greater variability in stroke patterns for a particular letter, both within and across participants, may be related to the fact that writing certain letters requires more violations of basic principles than writing other letters does. The final-Tsadi in Hebrew print is almost an exact left-right reversal of a lowercase y, and therefore presents similar challenges to y. As we have discussed in Chapter 4, letters containing oblique lines (such as y, k, and final-Tsadi) may present a particular challenge since avoiding a right-to-left movement might require a down-to-up movement, and vice versa. With the letter Tet there is again the issue of clashing principles, since a minor-stroke (the one at the top-right of the letter Tet: ט) should be produced after all major strokes, but the stroke order in Hebrew should move from right-to-left, and avoid pen lifts. Other cases in which any stroke pattern would necessarily violate a constraint present similar issues, and might therefore contribute both to our difficulty to model these letters and to the greater variance we see in the field of stroke patters for that letter.

An interesting question that arises from the above, is how shapes are chosen for a language. Primus (2004) has used OT to examine letter-shapes in the Roman alphabet. She concludes that the internal structure of letters in the Roman alphabet is “highly systematic in […] inner graphematic terms” (Primus, 2004, p. 269), and explainable with constraints ranked in strict domination. In fact, she argues that the Roman letter-shapes are the optimal shapes of all their alternatives, and that they can be distinguished from digits and punctuation marks using differences in domination relations. While Primus’ work is related and valuable to our enterprise, she only considered the base-shape of letters (as well as their phonological functions), but not features related to the handwritten production of the letter. For example, she observes that Roman letters tend to be open to the right (e.g., the letters C, G, or E), whereas digits tend to be open to the left (e.g., 3, 9). However, she does not comment on the direction in which such shapes ought to be produced, or on the way people actually produce them. Examining letter shapes in light of Primus' work could shed light not only on the shapes that made it into a language, but also on the degree of compatibility of new shapes with a given alphabet.

Future directions

Effector-independent motor plans

The work in this dissertation constitutes a foundation upon which new research can build. One of the exciting opportunities that stems directly from this work concerns the nature of abstract, or effector-independent motor-plans. Previous research (e.g., Ellis, 1988; Margolin, 1984) has proposed that letter shapes are represented at different levels, some effector-independent and others effector-specific. While many researchers discuss the existence of effector-independent motor plans, different researchers have characterized them quite differently, and often in very general or vague terms. For example, Wong, Haith, and Krakauer (2015) describe the abstract motor plans as an "optional process" involving "decisions about the shape of the trajectory to be produced in an effector-independent manner" (Wong, Haith, & Krakauer, 2015, p. 390). Saltzman (1979) describes the "effector-nonspecific level" as "an abstract specification of a goal in terms of desired relationships between objects [… in which] the spatial and temporal aspects of motion are expressed only in an abstract qualitative sense" (Saltzman, 1979, p. 98). Other researchers, such as Carter and Shapiro (1984), take a more compromising approach, describing the "generalized motor program" as "an abstract memory structure composed of [both] invariant and variant characteristics" (Carter & Shapiro, 1984, p. 788).

The "abstractness", or the "effector-independence" of a motor plan can be characterized in one of two ways. Under a narrow definition, a motor plan is effector-independent if it simply does not specify the particular muscles to be used to execute it. This is the definition we have adopted for the majority of this work. The motor plan in this case will specify the trajectory and order of movements (in our case – the direction and order of strokes to produce a letter), but those can be executed using any effector. We call this type of effector-independent motor plan a "fixed" plan. Under a broader definition, a motor plan is effector-independent if it is stated in a sufficiently abstract, or variable way that it can be easily applied to create *optimal* effector-specific programs, regardless of the effector. We name this latter type a "variant" plan.

A fixed effector-independent motor plan gives rise to a reasonable hypothesis, compatible with Merton (1972), and Raibert (1977), under which learned effector-independent motor plans detail all of the trajectories needed to produce a movement, but not the specific muscles needed to execute them. In our case, the motor plans would be established when learning how to write, and include the direction and order of production for each of the base strokes, as well as the trajectory and direction of transition strokes. Presumably, the abstract motor plans that writers develop after mastering writing are optimized, among other variables, for their preferred effector (i.e., their dominant writing hand), although the plans they are originally taught could be sub-optimal for left-handed writers. Once the motor plans are established for the dominant hand, they can be used to write with the non-dominant hand, even though they may be sub-optimal for this effector, or a writer could instead generate new effector-independent motor plans of the fixed type (i.e., motor plans that could be used by either effector, but are optimized for the non-dominant hand) to accommodate the change.

If no new motor plans are generated for the new effector, one would expect to see the same stroke patterns used by participants when writing with their non-dominant hand as they do when writing with their dominant hand. However, as pointed out by Wright (1990), among others, different effectors in the human body have such different geometries, that it would be "quite extraordinary", per Wright, to use the same motor plan, with the same trajectories, with a different effector. And indeed, a preliminary analysis of writing with the non-dominant hand that we have collected from our right and left-handed participants shows that participants do in fact change their stroke patterns when writing with their non-preferred effector.

The main difference we see between writing with the non-dominant hand compared to the dominant hand is a change in the direction of horizontal strokes (e.g., when writing with the right hand, participants produce them from left-to-right, and vice versa for the left hand). As we explained in Chapter 3, this result is expected, as abduction movement would dictate the opposite direction depending on the hand, and it could be more efficient than adduction movement (Bradshaw et al., 1990). The change in stroke patterns when switching effectors is evidence against the claims by Merton (1972) and Raibert (1977) that participants use the same fixed effector-independent motor plan regardless of the effector. Instead, these initial results might indicate that participants either generate a new fixed motor plan to accommodate the new effector (a possible modus operandi, albeit not a very cost-effective one), or the motor plan being used is of the variant type.

However, this second option, that the effector-independent motor plans are phrased entirely in “variant” terms (e.g., in terms of adduction/abduction movement), might face its own theoretical problems, raising the question of whether it is at all viable. To demonstrate this, let us consider what such a program would contain. It is immediately apparent that it would not be possible to specify the existence or the trajectory of transition strokes without knowing where each base stroke begins and ends. Consider for example the letter L. If the horizontal stroke begins where the vertical stroke ends (as in the prescribed production of L), there will not be a transition stroke between them. On the other hand, if, say, the horizontal stroke is produced after the vertical stroke, and from right-to-left, there would have to be a transition stroke from the end of the vertical stroke to the beginning of the horizontal one.

It is also not straightforward that the order of base-strokes could always be specified under these conditions. The exact order of strokes might depend on where a stroke begins and ends, in conjunction with other considerations. For example, a motor plan which aims to minimize the number of pen lifts (and in which top-down direction for vertical strokes is preferred to any effector-specific direction) might have a different ordering of the strokes in the letter ח (/xet/, Figure 17) depending on the effector. For right-handed production, an abduction movement will have stroke y produced from left-to-right, and to minimize pen lifts it would then specify the production of stroke z immediately following stroke y. On the other hand, left-handed production will have stroke y produced from right-to-left, and then stroke x produced immediately after it (Figure 17). Thus, if the direction of strokes is specified in terms of adduction/abduction movement, the order of strokes depends on other parameters (or other constraints) for which the program optimizes, and would not be able to be specified before an effector is chosen.

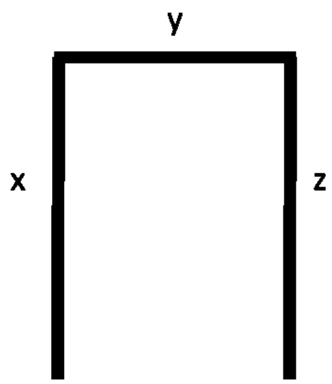


Figure 17. A representation of the three strokes comprising the letter ח. Assuming a preference for abduction movement, stroke y will be produced in a different direction with each hand. If we assume a preference for top-down strokes and minimizing pen lifts, not just the direction of stroke y, but the order of the strokes would have to change to accommodate a different effector.

Moreover, even the direction of production for individual base strokes cannot necessarily be specified under this hypothesis. Consider, once more, the uppercase letter L. While the vertical line will presumably be optimal when produced from the top-down using either effector, an abduction movement for the lower horizontal line will be left-to-right for the right hand, and right-to-left for the left hand. Assume then that the motor plan aims to optimize for the smallest number of pen lifts. For production with the right hand we might not have any conflict, but production with the left hand will face a dilemma – either lift the pen and produce an abduction movement, or do not lift the pen but produce an adduction movement.

An effector-independent plan in the variant sense we outlined above will only be able to specify the direction of motion under one of two conditions: Either a) the constraint specifying the direction of the horizontal stroke is ranked lower than all other constraints, in which case the motor-plan may very well be determined independently from the effector, or b) the variant constraint, specifying the direction of strokes in terms of adduction/abduction is the only constraint distinguishing between two candidates. The latter option can be exemplified by the allograph of Z with the crossing line in the middle: Ƶ. While other constraints may dictate the direction of production for the top and bottom horizontal strokes, as well as the order of strokes, only the constraint concerning direction of strokes will apply to the crossing line. Therefore, in this case there can be a fully-specified effector-independent motor program with a variant parameter to be filled out only when choosing the effector.

But this is not usually the case. What happens if the variant constraint is ranked higher than other constraints? One possibility is that there is no "effector-independent motor plan" at all, but instead there is a ranking of constraints, which are phrased in terms of adduction/abduction, and the effector-specific motor programs are only calculated on the basis of the constraint ranking once a particular effector is chosen. This hypothesis involves a mix of the two senses of effector-independent we outlined above: The constraints are variant and effector-independent, but an optimal program, including the exact stroke pattern, is computed only when an effector is chosen.

Under a strong version of this last hypothesis, we should expect to see the same constraints, and the same ranking, applying to both the dominant and the non-dominant hand of a given participant, with only the direction of adduction/abduction movement realized differently according to the selected effector. If we see a different ranking of constraints for the non-dominant hand, or if we are unable to find a ranking of constraints for the non-dominant hand, then this strong version of the hypothesis could be rejected. A softer (and perhaps more realistic) version of this hypothesis might suggest that the optimal stroke patterns for a given effector will need to be computed based on constraints tailored to a specific effector. In this latter case, the new stroke patterns may be learned, but possibly not completely or immediately. We will thus see stroke patterns which start out similar to the over-learned patterns used by the dominant hand millions of times, but gradually (and possibly not at the same rate for all letters) shift into stroke patterns that are more convenient for the non-dominant hand.

This last pattern of performance would also be compatible with the hypothesis of abstract or effector-independent motor-plans in the fixed sense (i.e., the non-muscle specific plans). The stroke patterns we expect in this case will be much less clear. We might not be able to model the writing at all (seeing as some letters will be produced following a stroke pattern appropriate for the dominant hand, and others following a stroke pattern appropriate for the non-dominant hand), and we might see a difference in stroke patterns between the beginning and end of a writing session. While the modeling of participants' writing when using their non-dominant hand is outside the scope of this dissertation, we have collected data from both right- and left-handed English-speakers writing with their dominant and non-dominant hand, which we have started analyzing in the hope of learning more about the nature and existence of effector-independent motor plans.

Effects of context, culture, and instruction

People not only have to develop a handwriting style that specifies how to write each shape, but in fact everyone has been taught to write letters a certain way. The starting point for proficient writers is the prescribed stroke patterns they were taught in kindergarten. Then they might diverge from those stroke patterns, and they might change how they write some letters but not others. What motivates people's stroke patterns is not a clean collection of inherently-motivated principles, but an amalgam of the principles, the instruction they received on writing the letters, and other cultural and contextual effects.

In a recent article for Quartz (Ha & Sonnad, 2017), the writers analyze the data from over 119,000 unique drawings of circles by people all over the world using Google's "Quick, Draw!", and note that the direction of circle production is significantly correlated with country of origin. People in the US, for example, produce circles clockwise about 14% of the time, compared with people from Japan who produce them clockwise 80% of the time. The authors of the Quartz article did not analyze some of the factors that we have, such as the point of origin on the circle's circumference, or people's preferred writing-hand. Furthermore, from Ha & Sonnad's results (2017), like ours, it is difficult to tease apart the effects of instruction, culture, and unique aspects of the language (e.g., our results show that Hebrew-speakers tend to produce circles clockwise more often than English-speakers, and this might be related to the fact that Hebrew is written and read from right-to-left). Other research has looked at the correlation in time between changes in stroke patterns and writing instruction. For example, Goodnow and Levine (1973), who looked at changes in preferences with development, note a shift in childrens' preferences regarding stroke order around the time they learn how to write in school.

One way to investigate this without looking at the writing of children (which may present other challenges) is to look at the differences in stroke patterns between participants who have been taught how to produce certain shapes and participants who have not been taught those shapes. We have collected data from Hebrew speakers writing English and English-speakers copying Hebrew. A thorough analysis of these data is planned for a future paper, and could shed light on the effect of language context. Since the Hebrew-speakers all read and write English, we can compare their writing of two different languages that they have been taught how to write, presumably isolating such factors as the direction of reading and writing from the influence of cultural or individual characteristics. We can also use these data to compare the writing of Hebrew by people who speak the language, and were instructed on how to write it, from people who have not been told how to produce these shapes.

Finally, we note that we were unable to model any Hebrew-speakers' writing of Hebrew print, despite being able to model the prescribed stroke patterns. Hebrew speakers who live in Israel (as do all of our Hebrew participants), are exposed to print letters every day, and in fact they are probably exposed to them to a significantly greater extent than to Hebrew script. However, after 1st grade, they hardly ever produce the print form in writing, and are only exposed to it in reading. We can therefore hypothesize, in line with other research (e.g., Kersey & James, 2013; Wiley et al., 2016; Wong et al., 2018), that handwriting experience is crucial for a stable, detailed representation of the letter. That detailed representation contains not just information about the shape, but also about the application of constraints to the strokes, and perhaps, about the constraints' ranking.

Conclusions

This dissertation describes four experimental analyses, using the framework of ranked violable constraints: 1) Modeling of prescribed writing in English (upper and lowercase), Hebrew print, and Hebrew script, using a novel application of Optimality Theory (OT); 2) modeling of participant writing in English lowercase and Hebrew script using OT; 3) modeling of right- and left-handed participants' writing in English lowercase using OT; and 4) modeling prescribed writing of English and Hebrew as well as RH and LH participants' writing in English using Harmonic Grammar.

In this dissertation we have established that we can account for both prescribed and actual writing, in both English and Hebrew, and for both right- and left-handed individuals within the framework of violable ranked constraints. We then used the modeling to further our understanding of some key differences between writing systems and styles, and between the writing of different individuals. Using the OT framework, we were able to illuminate complex and opaque relationships between rules and provide answers to previously unresolved debates (e.g., regarding the direction of production of letters in Hebrew). We were able to highlight some of the ways in which prescribed writing is different from actual writing (e.g., when pen lifts occur), to illustrate the differences in individual writing styles using direct contrasts of constraints (e.g., circles clockwise vs. counter-clockwise), and to settle the long-standing debate regarding writing direction vs. control of motion in right-handed Hebrew writers and in left-handed English writers.

The work we have presented demonstrates the viability of using ranked violable constraints to account for stroke-patterns in handwritten production, and constitutes the foundations for the usage of the OT framework for the study of handwriting. It opens additional avenues of research concerning some of the fundamental cognitive mechanisms underlying handwriting (e.g., the existence and nature of effector-independent motor plans), and could be useful for other fields of cognition (e.g., motor function in general). We conclude that our particular implementation of both OT and HG is useful to this domain, and helps shed light not only on the reasons behind people's choice of certain stroke patterns, but on the differences between people, between languages, and between prescribed and actual writing.

Appendix A

Table 20. The constraints used in our analysis of prescribed and participant writing in English and Hebrew. Variant Definition details the exact calculation that was done for each version of constraint.

CATEGORY	CONSTRAINT	VARIANT
STROKE DIRECTION	no down-to-up base strokes	1 violation for each base stroke with end y > (above) start y
	no right-to-left base strokes	1 violation for each base stroke with end x < (left of) start x
	no left-to-right base strokes	1 violation for each base stroke with end x > (right of) start x
	no initial down-to-up base strokes	1 violation if first base stroke ends higher than it starts
	no down-to-up curved base strokes	1 violation if curved stroke ends higher than it started
	no down-to-up vertical base strokes	1 violation if base stroke is straight, start x = end x and end y > (above) start y
STROKE SEQUENCE	no down-to-up sequencing of base strokes v1	1 violation for each pair of base strokes in which entire later stroke is above any part of the earlier stroke
	no down-to-up sequencing of base strokes v2	1 violation for each pair of base strokes in which any part of the later stroke is above any part of the earlier stroke
	no right-to-left sequencing of base strokes v1	1 violation for each pair of base strokes in which any part of the later stroke is left of any part of the earlier stroke
	no right-to-left sequencing of base strokes v2	1 violation for each pair of base strokes in which the entire later stroke is left of any part of the earlier stroke

	no left-to-right sequencing of base strokes	1 violation for each pair of base strokes in which later stroke is entirely right of earlier stroke
	minor base strokes after major base strokes	1 violation for each pair of base strokes in which the earlier is a minor stroke that is not continued (i.e., where the next stroke doesn't start at its end) and the latter is a major stroke
	continue to adjacent base stroke without lifting the pen	1 violation if base stroke ends at point corresponding to terminal point of a not-yet-produced base stroke but does not continue to that stroke
	no high-precision intersections	1 violation if base stroke must end at, touch or pass through a specific point on a previously-produced base stroke
START POSITION	start at leftmost possible start point	1 violation if starting point of first base stroke is not leftmost possible base stroke starting point
	start at rightmost possible start point	1 violation if starting point of first base stroke is not rightmost possible base stroke starting point
	start with stroke containing point closest to upper left corner	1 violation if first base stroke does not contain point closest to upper left corner
	start with stroke containing point closest to upper right corner	1 violation if first base stroke does not contain point closest to upper right corner
	first base stroke starts between baseline & x line v1	1 violation if first base stroke starts below base line or above x line
	first base stroke starts between baseline & x line v2	1 violation for any base stroke starting in descender area
TRANSITION STROKES	no pen lifts	1 violation for each pen up transition stroke
	no transition strokes	1 violation for each transition stroke
	pen up on all transition strokes	1 violation for each pen down transition stroke

	pen up on transition strokes not in base-shape	1 violation for each transition stroke with one or more points that do not coincide with base stroke points
	transition stroke goes to closer end of next stroke	1 violation for any transition stroke that goes to farther end of next stroke
CURVES	curves continue prior motion direction	1 violation for any closed curve that does not continue the motion direction of a prior straight stroke
	curves counter-clockwise	1 violation for any closed curve in clockwise direction
	curves clockwise	1 violation for each curved base stroke with a counterclockwise direction
	closed curve start position determines motion direction	if start is left of center 1 violation if counter-clockwise; if start is right of center 1 violation if clockwise

Appendix B

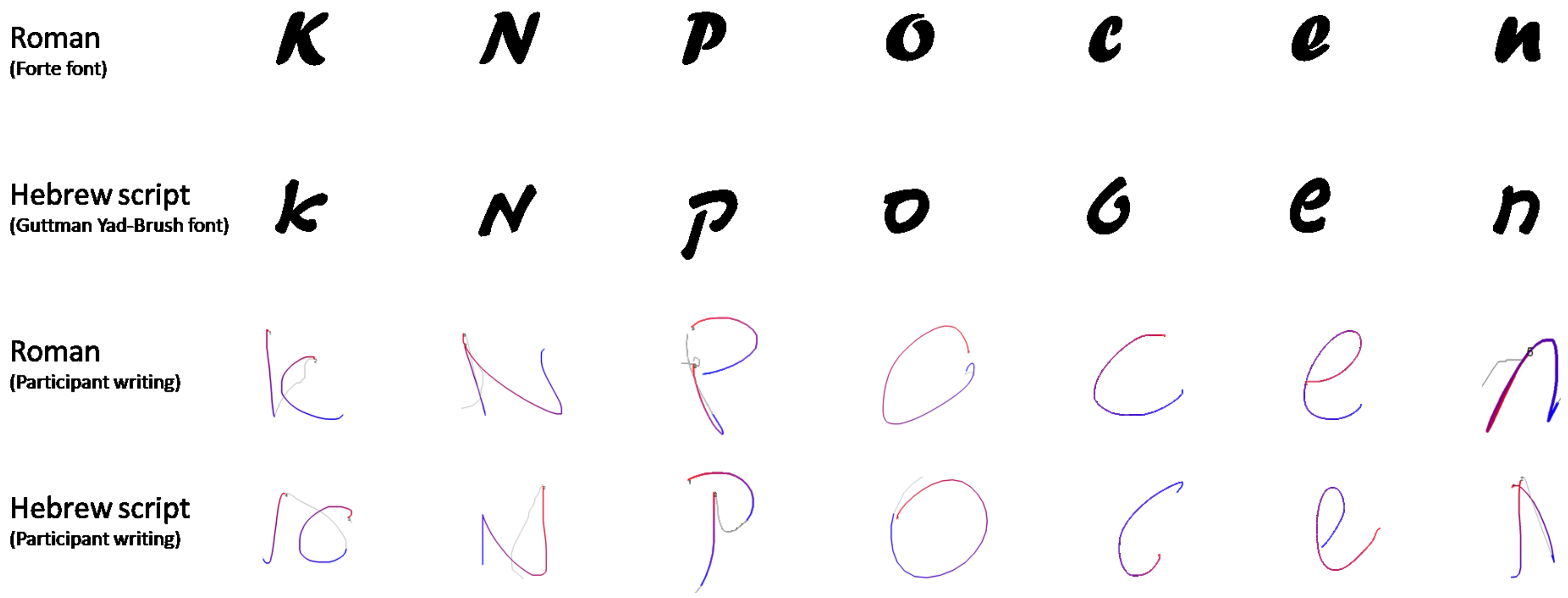


Figure 18. Letters that share the same general shape in Roman and Hebrew script. Roman (left-to-right): K, N, P, o, c, e, n. Hebrew (left-to-right): Aleph, Mem, Qof, Samech, Tet, Shin, Het. The top two rows show the letter shape in a computer-printed handwriting font (Forte for Roman and Guttman Yad-Brush for Hebrew). The bottom two rows show the letter shape as produced by participants when writing English (top) and Hebrew script (bottom).

References

Amenomori, M., Kono, A., Fournier, J. S., & Winer, G. A. (1997). A cross-cultural developmental study of directional asymmetries in circle drawing. *Journal of Cross-Cultural Psychology*, *28*(6), 730–742.

Berninger, V. W., Abbott, R. D., Abbott, S. P., Graham, S., & Richards, T. (2002). Writing and Reading: Connections Between Language by Hand and Language by Eye. *Journal of Learning Disabilities*, *35*(1), 39–56. https://doi.org/10.1177/002221940203500104

Boersma, P., Dekkers, J., & van de Weijer, J. (2000). Introduction - Optimality Theory: Phonology, syntax, and acquisition. In J. Dekkers, F. van der Leeuw, & J. van de Weijer (Eds.), *Optimality Theory: Phonology, Syntax, and Acquisition* (pp. 1–44). Oxford: Oxford University Press.

Bradshaw, J. L., Bradshaw, J. A., & Nettleton, N. C. (1990). Abduction, adduction and hand differences in simple and serial movements. *Neuropsychologia*, *28*(9), 917–931. https://doi.org/10.1016/0028-3932(90)90108-Z

Brasoveanu, A., & Prince, A. (2005). Ranking and necessity: The Fusional Reduction algorithm. *Part I: The Fusional Reduction Algorithm. [ROA-794]*. Retrieved from http://people.ucsc.edu/~abrsvn/ROA_794_Brasoveanu_Prince.pdf

Buchwald, A., & Rapp, B. (2004). Rethinking the graphemic buffer? *Brain and Language*, *91*(1), 100–101. https://doi.org/10.1016/j.bandl.2004.06.052

Caramazza, A., & Miceli, G. (1990). The structure of graphemic representations. *Cognition*, *37*, 243–297.

Carter, M. C., & Shapiro, D. C. (1984). Control of sequential movements: evidence for generalized motor programs. *Journal of Neurophysiology*, *52*(5), 787–796. https://doi.org/10.1152/jn.1984.52.5.787

Corey, D. M., Hurley, M. M., & Foundas, A. L. (2001). Right and left handedness defined, *14*(3), 144–152.

De Agostini, M., & Chokron, S. (2002). The Influence of handedness on profile and line drawing directionality in children, young, and older normal adults. *Brain and Cognition*, *48*(2–3), 333–365.

Edelman, S., & Flash, T. (1987). A model of handwriting. *Biological Cybernetics*, *57*(1–2), 25–36. https://doi.org/10.1007/BF00318713

Ellis, A. W. (1988). Normal writing processes and peripheral acquired dysgraphia. *Language and Cognitive Processes*, *3*(2), 99–127.

Ellis, A. W. (1989). *Reading, Writing and Dyslexia*. Hove, UK: LEA Publishers.

Fischer-Baum, S., McCloskey, M., & Rapp, B. (2010). Representation of letter position in spelling: Evidence from acquired dysgraphia. *Cognition*, *115*(3), 466–490. https://doi.org/10.1016/j.cognition.2010.03.013

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*(4), 365–387. https://doi.org/10.2307/356600

Goodnow, J. J., Friedman, S. L., Bernbaum, M., & Brauch Lehman, E. (1973). Direction and sequence in copying: The effect of learning to write in English and Hebrew. *Journal of Cross-Cultural Psychology*, *4*(3), 263–282.

Goodnow, J. J., & Levine, R. A. (1973). “The Grammar of Action”: Sequence and syntax in children’s copying. *Cognitive Psychology*, *4*(1), 82–98. https://doi.org/10.1016/0010-0285(73)90005-4

Guinet, E., & Kandel, S. (2010). Ductus: A software package for the study of handwriting production. *Behavior Research Methods*, *42*(1), 326–332. https://doi.org/10.3758/BRM.42.1.326

Ha, T.-H., & Sonnad, N. (2017). How do you draw a circle? We analyzed 100,000 drawings to show how culture shapes our instincts. Retrieved from https://qz.com/994486/the-way-you-draw-circles-says-a-lot-about-you/

Hardyck, C., & Petrinovich, L. F. (1977). Left-handedness. *Psychological Bulletin*, *84*(3), 385–404.

Hillis, A. E., & Caramazza, A. (1989). The Graphemic Buffer and attentional mechanisms. *Brain and Language*, *36*(2), 208–235. https://doi.org/10.1016/0093-934X(89)90062-X

Hillis, A. E., & Caramazza, A. (1991). Mechanisms for accessing lexical representations for output: Evidence from a category-specific semantic deficit. *Brain and Language*, *40*(1), 106–144. https://doi.org/10.1016/0093-934X(91)90119-L

Hillis, A. E., & Tippett, D. C. (2014). Stroke Recovery: Surprising Influences and Residual Consequences. *Advances in Medicine*, *2014*, 1–10. https://doi.org/10.1155/2014/378263

Hollerbach, J. M. (1980). *An oscillatory theory of handwriting*. Massachusetts Institute of Technology, Boston, MA.

James, K. H., & Engelhardt, L. (2012). The effects of handwriting experience on functional brain development in pre-literate children. *Trends in Neuroscience and Education*, *1*(1), 32–42. https://doi.org/10.1016/j.tine.2012.08.001

James, K. H., & Gauthier, I. (2009). When writing impairs reading: Letter perception's susceptibility to motor interference. *Journal of Experimental Psychology: General*, *138*(3), 416–431. https://doi.org/10.1037/a0015836

Kandel, S., & Perret, C. (2015). How do movements to produce letters become automatic during writing acquisition? Investigating the development of motor anticipation. *International Journal of Behavioral Development*, *39*(2), 113–120. https://doi.org/10.1177/0165025414557532

Kandel, Sonia, Peereman, R., & Ghimenton, A. (2013). Further evidence for the interaction of central and peripheral processes: The impact of double letters in writing English words. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00729

Kandel, Sonia, & Spinelli, E. (2010). Processing complex graphemes in handwriting production. *Memory & Cognition*, *38*(6), 762–770. https://doi.org/10.3758/MC.38.6.762

Kersey, A. J., & James, K. H. (2013). Brain activation patterns resulting from learning letter forms through active self-production and passive observation in young children. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00567

Lake, B. M., Salakhutdinov, R., Gross, J., & Tenenbaum, J. B. (2011). One shot learning of simple visual concepts. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (Vol. 172, p. 2). Retrieved from http://palm.mindmodeling.org/cogsci2011/papers/0601/paper0601.pdf

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2012). Concept learning as motor program induction: A large-scale empirical study. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 659–664). Retrieved from http://www.utstat.toronto.edu/~rsalakhu/papers/LakeEtAl2012CogSci.pdf

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2013). One-shot learning by inverting a compositional causal process. In *Advances in neural information processing systems* (pp. 2526–2534). Retrieved from http://papers.nips.cc/paper/5128-one-shot-learning-by-inverting-a-compositional-causal-process

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, *350*(6266), 1332–1338.

Law, N., Ki, W. W., Chung, A. L. S., Ko, P. Y., & Lam, H. C. (1998). Children’s stroke sequence errors in writing Chinese characters. In C. K. Leong & K. Tamaoka (Eds.), *Cognitive Processing of the*

Chinese and the Japanese Languages (Vol. 14, pp. 113–138). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-015-9161-4_7

Legendre, G., Miyata, Y., & Smolensky, P. (1990a). *Harmonic Grammar - a formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations* (ICS Technical Report No. 90–5). Colorado: University of Colorado at Boulder.

Legendre, G., Miyata, Y., & Smolensky, P. (1990b). *Can connectionism contribute to syntax? Harmonic Grammar, with an application* (ICS Technical Report No. 90–12). Colorado: University of Colorado at Boulder.

Margolin, D. I. (1984). The neuropsychology of writing and spelling: Semantic, phonological, motor, and perceptual processes. *The Quarterly Journal of Experimental Psychology Section A*, *36*(3), 459–489. https://doi.org/10.1080/14640748408402172

McCloskey, M., Badecker, W., Goodman-schulman, R. A., & Aliminosa, D. (1994). The structure of graphemic representations in spelling: Evidence from a case of acquired dysgraphia. *Cognitive Neuropsychology*, *11*(3), 341–392. https://doi.org/10.1080/02643299408251979

McCloskey, M., & Rapp, B. (2017). Developmental dysgraphia: An overview and framework for research. *Cognitive Neuropsychology*, *34*(3–4), 65–82. https://doi.org/10.1080/02643294.2017.1369016

McCloskey, M., Reilhac, C., & Schubert, T. (2018). A deficit in post-graphemic writing processes: Evidence for a graphomotor buffer. *Cognitive Neuropsychology*, *35*(8), 430–457. https://doi.org/10.1080/02643294.2018.1546683

Mckeever, W. F. (1979). Handwriting posture in left-handers: Sex, familial sinistrality, and language laterality correlates. *Neuropsychologia*, *17*, 429–444.

Merton, P. A. (1972). How we control the contraction of our muscles. *Scientific American*, *226*(5), 30–37.

Meulenbroek, R. G. J., & van Galen, G. P. (1989). The production of connectionist strokes in cursive writing: Developing co-articulation in 8 to 12-year-old children. In R. Plamondon, C. Y. Suen, & M. Simner (Eds.), *Computer recognition and human production of handwriting* (pp. 273–286). Singapore: World Scientific.

Meulenbroek, R. G. J., Vinter, A., & Mounoud, P. (1993). Development of the start-rotation principle in circle production. *British Journal of Developmental Psychology*, *11*(3), 307–320. https://doi.org/10.1111/j.2044-835X.1993.tb00605.x

Miozzo, M., & De Bastiani, P. (2002). The Organization of Letter-Form Representations in Written Spelling: Evidence from Acquired Dysgraphia. *Brain and Language*, *80*(3), 366–392. https://doi.org/10.1006/brln.2001.2598

Nihei, Y. (1983). Developmental changes in covert principles for the organization of strokes in drawing and handwriting. *Acta Psychologica*, *54*, 221–232.

Ninio, A., & Lieblich, A. (1976). The Grammar of Action: "Phrase structure" in children's copying. *Child Development*, *47*(3), 846. https://doi.org/10.2307/1128205

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Palmis, S., Danna, J., Velay, J.-L., & Longcamp, M. (2017). Motor control of handwriting in the developing brain: A review. *Cognitive Neuropsychology*, *34*(3–4), 187–204. https://doi.org/10.1080/02643294.2017.1367654

Pater, J. (2009). Weighted Constraints in Generative Linguistics. *Cognitive Science*, *33*(6), 999–1035. https://doi.org/10.1111/j.1551-6709.2009.01047.x

Peters, M., & McGrory, J. (1987). The writing performance of inverted and noninverted right- and left-handers. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *41*(1), 20–32. https://doi.org/10.1037/h0084146

Picard, D. (2011). Impact of manual preference on directionality in children's drawings. *Laterality: Asymmetries of Body, Brain and Cognition*, *16*(1), 24–34. https://doi.org/10.1080/13576500903140610

Potts, C., Pater, J., Jesney, K., Bhatt, R., & Becker, M. (2010). Harmonic Grammar with linear programming: from linear systems to linguistic typology. *Phonology*, *27*(01), 77–117. https://doi.org/10.1017/S0952675710000047

Primus, B. (2004). A featural analysis of the Modern Roman Alphabet. *Written Language & Literacy*, *7*(2), 235–274.

Prince, A. (2002). Anything goes. In T. Honma, M. Okazaki, T. Tabata, & S. ichi Tanaka (Eds.), *New century of phonology and phonological theory* (pp. 66–90). Tokyo: Kaitakusha.

Prince, A., & Smolensky, P. (1993/2004). *Optimality Theory: Constraint interaction in generative grammar.* Technical report, Rutgers University and University of Colorado, Boulder, 1993. Rutgers Optimality Archive 537, 2002. Revised version published by Blackwell, 2004.

Purcell, J. J., Jiang, X., & Eden, G. F. (2017). Shared orthographic neuronal representations for spelling and reading. *NeuroImage*, *147*, 554–567. https://doi.org/10.1016/j.neuroimage.2016.12.054

Purcell, J. J., Turkeltaub, P. E., Eden, G. F., & Rapp, B. (2011). Examining the Central and Peripheral Processes of Written Word Production Through Meta-Analysis. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00239

Raibert, M. H. (1977). *Motor control and learning by the state space model* (PhD). Massachusetts Institute of Technology, Cambridge, MA.

Rapp, B., & Caramazza, A. (1997). From graphemes to abstract letter shapes: Levels of representation in written spelling. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(4), 1130.

Rapp, B., Purcell, J., Hillis, A. E., Capasso, R., & Miceli, G. (2016). Neural bases of orthographic long-term memory and working memory in dysgraphia. *Brain*, *139*(2), 588–604. https://doi.org/10.1093/brain/awv348

Sáenz-Badillos, A. (1993). *A history of the Hebrew language*. Cambridge, UK: Cambridge University Press.

Saltzman, E. (1979). Levels of sensorimotor representation. *Journal of Mathematical Psychology*, *20*(2), 91–163. https://doi.org/10.1016/0022-2496(79)90020-8

Simner, M. L. (1984). The grammar of action and reversal errors in children’s printing. *Developmental Psychology*, *20*(1), 136.

Taguchi, M., & Noma, Y. (2005). Relationship between directionality and orientation in drawings by young children and adults. *Perceptual and Motor Skills*, *101*, 90–94.

Tainturier, M.-J., & Rapp, B. (2001). The spelling process. In B. Rapp (Ed.), *The Handbook of Cognitive Neuropsychology: What Deficits Reveal about the Human Mind* (pp. 263–289). New York, NY: Psychology Press.

Tesar, B., & Smolensky, P. (1998). Learnability in Optimality Theory. *Linguistic Inquiry*, *29*(2), 229–268.

Thomassen, A. J. W. M., Meulenbroek, R. G. J., & Tibosch, H. J. C. M. (1991). Latencies and kinematics reflect graphic production rules. *Human Movement Science*, *10*(2–3), 271–289. https://doi.org/10.1016/0167-9457(91)90007-K

van Galen, G. P. (1991). Handwriting: Issues for a psychomotor theory. *Human Movement Science*, *10*(2–3), 165–191. https://doi.org/10.1016/0167-9457(91)90003-G

Van Sommers, P. (1984). *Drawing and cognition: Descriptive and experimental studies of graphic production processes*. Cambridge, UK: Cambridge University Press.

Vrandečić, D. (2012, November 4). How often is which letter. Retrieved from http://simia.net/letters/

Wiley, R. W., Wilson, C., & Rapp, B. (2016). The effects of alphabet and expertise on letter perception. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(8), 1186–1203. https://doi.org/10.1037/xhp0000213

Wing, A. M. (1979). Variability in handwritten characters. *Visible Language*, *13*(3), 283–298.

Wong, A. L., Haith, A. M., & Krakauer, J. W. (2015). Motor Planning. *The Neuroscientist*, *21*(4), 385–398. https://doi.org/10.1177/1073858414541484

Wong, K., Wadee, F., Ellenblum, G., & McCloskey, M. (2018). The devil's in the g-tails: Deficient letter-shape knowledge and awareness despite massive visual experience. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(9), 1324–1335. https://doi.org/10.1037/xhp0000532

Wright, C. E. (1990). Generalized motor programs: Reexamining claims of effector independence in writing. In *Attention and Performance XIII: Motor Representation and Control* (pp. 294–320). Hillsdale, NJ: Lawrence Erlbaum Associates.

Zesiger, P., Mounoud, P., & Hauert, C.-A. (1993). Effects of lexicality and trigram frequency on handwriting production in children and adults. *Acta Psychologica*, *82*(1–3), 353–365. https://doi.org/10.1016/0001-6918(93)90020-R

Curriculum vitae

Gali Ellenblum was born in Jerusalem, Israel, on October 30th, 1986. She studied in the Adi Lautman Interdisciplinary Program for Outstanding Students at Tel Aviv University, and in 2013 completed her MA degree in Cognitive Psychology under the supervision of Dr. Naama Friedmann. At Tel Aviv University she taught Statistics for Psychologists, and worked in Dr. Friedmann's Brain and Language lab and in Dr. Talma Hendler's Brain and Emotional Experience lab. In September of 2013 she joined the Cognitive Science Department at Johns Hopkins University. She completed an MA thesis under the supervision of Dr. Brenda Rapp, examining the resting-state functional connectivity of orthographic processing networks. She completed her PhD under the mentorship of Dr. Michael McCloskey, on cognitive and computational modeling of handwriting.